

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 06:19:20 ; Search time 6510 Seconds
(without alignments)
7700.013 Million cell updates/sec

Title: US-09-597-920B-1

Perfect score: 1060

Sequence: 1 gactctgcctcctgaggggccc.....aaaaaaaaaaaaaaaaaaaa 1060

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_loc:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sls:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1060	100.0	1060	9	AF036905 Homo sapi
2	1027.4	96.9	1616	9	AJ223280 Homo sapi
3	1013.4	95.6	1472	9	BC011563 Homo sapi
4	949.2	89.5	1224	6	BD107841 36 human
5	932	87.9	1460	9	AF036906 Homo sapi
6	479.6	45.2	2443	9	AK093815 Homo sapi
7	416.8	39.3	528	6	AR380020 Sequence
8	413.8	39.0	1260	10	BC052340 Mus muscu
9	400.4	37.8	1260	10	AF036907 Mus muscu
10	396	37.4	1188	10	AJ001184 Rattus no
11	395.4	37.3	1232	10	BC013337 Mus muscu
12	265	25.0	131839	9	AC112166 Homo sapi
13	265	25.0	163424	9	AC109460 Homo sapi
14	265	25.0	187788	2	AC138925 Homo sapi
15	265	25.0	219646	2	AC139247 Homo sapi
16	205.2	19.4	400	6	CQ664157 Sequence
17	200.6	18.9	201	11	BV204291 sqm21471
18	100.6	9.5	182032	2	AC150435 Atelelix
19	67.4	6.4	125020	9	AF429315 Homo sapi

c 20	59.4	5.6	125020	9	AF429315 Homo sapi
c 21	56.2	5.3	140031	2	AC121686 Rattus no
c 22	56.2	5.3	231198	2	AC126892 Rattus no
23	54.6	5.2	6307	6	AX818044 Sequence
24	54.6	5.2	6307	6	AX826982 Sequence
25	54.6	5.2	6307	10	MMU438435 Mus muscu
c 26	54.6	5.2	217912	10	AC125322 Mus muscu
c 27	53.8	5.1	101451	2	AC138105 Mus muscu
28	53.2	5.0	2000	6	AX555393 Sequence
c 29	53	5.0	189190	2	AC129536 Mus muscu
c 30	52.6	5.0	2030	9	BC041637 Homo sapi
31	51.4	4.8	7218	6	166494 Sequence 14
c 32	51.2	4.8	101451	2	AC138105 Mus muscu
33	51	4.8	218780	2	AC115552 Rattus no
c 34	48.6	4.6	179886	2	AC140072 Mus muscu
c 35	48.6	4.6	219568	10	AL844840 Mouse DNA
36	48.4	4.6	162738	9	AC083812 Homo sapi
37	48.4	4.6	241432	10	AL589661 Mouse DNA
38	48	4.5	2285	9	AK000291 Homo sapi
39	48	4.5	231462	2	AC121752 Rattus no
40	47.6	4.5	194018	10	AC124686 Mus muscu
c 41	47.4	4.5	1610	8	AK109707 Oryza sat
42	47.4	4.5	3196	6	BD232172 Novel sec
43	47.4	4.5	123954	8	AC123518 Oryza sat
44	47.4	4.5	139943	2	AC119190 Mus muscu
45	47.4	4.5	162311	8	AC145321 Oryza sat

ALIGNMENTS

RESULT 1	AF036905	1060 bp	mRNA	linear	PRI 06-MAR-1998
LOCUS	AF036905	Homo sapiens linker for activation of T cells (LAT) mRNA, complete cds.			
DEFINITION	AF036905	Homo sapiens linker for activation of T cells (LAT) mRNA, complete cds.			
ACCESSION	AF036905	AF036905			
VERSION	AF036905.1	GI:2828023			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Zhang,W., Sloan-Lancaster,J., Klitchen,J., Tribble,R.P. and Samelson,L.E.				
TITLE	LAT: the ZAP-70 tyrosine kinase substrate that links T cell receptor to cellular activation				
JOURNAL	Cell 92 (1), 83-92 (1998)				
MEDLINE	98149346				
PUBMED	9489702				
REFERENCE	2 (bases 1 to 1060)				
AUTHORS	Zhang,W., Sloan-Lancaster,J., Klitchen,J., Tribble,R.P. and Samelson,L.E.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-DEC-1997) Cell Biology and Metabolism Branch, National Institute of Child Health and Development, National Institute of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA				
COMMENT	LAT is a highly tyrosine phosphorylated protein, previously described as p36-38, and it associates with many signaling molecules, such as Grb2, PLC-gamma1, PI-3 kinase, cbl, Vav, and SHP-76, either directly or indirectly upon T cell activation. It is a potential type III transmembrane protein.				
FEATURES	Location/Qualifiers				
Source	1..1060				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/cell_line="Jurkat T cells"				
	1..1060				
gene	/gene="LAT"				
	/note="linker for activation of T cells"				
CDS	58..759				

```

/gene="LAT"
/note="tyrosine kinase substrate"
/codon_start=1
/product="LAT"
/protein_id="AAC39636.1"
/db_xref="GI:2828024"
/translation="MEBALTVCVLGILLPIILMLMALCVHCHRLPGSYDSTSSDSL
YPRGIQKRPHTVAWPWPAPVPTVSYPPLSQPDLPIPRSPQLGSHRTPSSRSD
GANSVASYENBEPACEDADEDDYHNHGYLVLPDSTPATSTPAASAPALSTPGIRD
SAFSMESIDYVNVVPESGESABASLDGSRVNVVSQELHGAATKTEPALSSQAEV
EEBAPPYENLQELN"

```

ORIGIN

Query Match 100.0%; Score 1060; DB 9; Length 1060;
 Best Local Similarity 100.0%; Pred. No. 7.9e-257;
 Matches 1060; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GACTCTGCTTGAAGGGGCTTAAGGGGTGACGACGCTGCTCCAGCTCCCTGCAATG 60
DB 1 GACTCTGCTTGAAGGGGCTTAAGGGGTGACGACGCTGCTCCAGCTCCCTGCAATG 60
QY 61 GAGAGGGCCATCTGGTCCCTGGGCTGCTGGGGGCTCCTGCTGTCGCCCATCTGGCCATG 120
DB 61 GAGAGGGCCATCTGGTCCCTGGGCTGCTGGGGGCTCCTGCTGTCGCCCATCTGGCCATG 120
QY 121 TTGATGGCACTGTGTGTGTCACAGTGCACAGTGCAGGCTCTTACGACAGCAATCTCA 180
DB 121 TTGATGGCACTGTGTGTGTCACAGTGCACAGTGCAGGCTCTTACGACAGCAATCTCA 180
QY 181 GATAGTTTGTATCCAAAGGGGATCCAGTTCAAAGGGCTTCAACAGGTTGGCCCTGCGCA 240
DB 181 GATAGTTTGTATCCAAAGGGGATCCAGTTCAAAGGGCTTCAACAGGTTGGCCCTGCGCA 240
QY 241 CCTGCTTACCCACCTGTACCTCTTACCTCTTACCTCTTACCTCTTACCTCTTACCTCT 300
DB 241 CCTGCTTACCCACCTGTACCTCTTACCTCTTACCTCTTACCTCTTACCTCTTACCTCT 300
QY 301 CCAAGATCCCGGAGCCCTTGGGGGCTCCAGCGGAGCCCATCTTCCCGGCGGATTTCT 360
DB 301 CCAAGATCCCGGAGCCCTTGGGGGCTCCAGCGGAGCCCATCTTCCCGGCGGATTTCT 360
QY 361 GATGCTGCAACAGTGTGGGAGCTACGAAACGAGAAACGAGCTGTGAGAGTGCAGAT 420
DB 361 GATGCTGCAACAGTGTGGGAGCTACGAAACGAGAAACGAGCTGTGAGAGTGCAGAT 420
QY 421 GAGATGAGAGCACTATCAACACCAAGCTTACCTGATGATGCTTCTGACAGCACCCCG 480
DB 421 GAGATGAGAGCACTATCAACACCAAGCTTACCTGATGATGCTTCTGACAGCACCCCG 480
QY 481 GCCACTAGCACTGTGCCCCCATCACTGCTCTGACAGCACCCCTGGCATCCGAGACGT 540
DB 481 GCCACTAGCACTGTGCCCCCATCACTGCTCTGACAGCACCCCTGGCATCCGAGACGT 540
QY 541 GCTCTTTCATGAGTCCATTGATGATTAAGTAAAGTTCGAGAGAGCGGGAGAGCCCA 600
DB 541 GCTCTTTCATGAGTCCATTGATGATTAAGTAAAGTTCGAGAGAGCGGGAGAGCCCA 600
QY 541 GCTCTTTCATGAGTCCATTGATGATTAAGTAAAGTTCGAGAGAGCGGGAGAGCCCA 600
DB 541 GCTCTTTCATGAGTCCATTGATGATTAAGTAAAGTTCGAGAGAGCGGGAGAGCCCA 600
QY 601 GAAAGGTCTCGAGTGGAGCGGGAGTATGATGATGATGATGATGATGATGATGATGATG 660
DB 601 GAAAGGTCTCGAGTGGAGCGGGAGTATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 GCGGCTAGAGCTGAGCTGCGCCCTGAGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 661 GCGGCTAGAGCTGAGCTGCGCCCTGAGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 GGGGCTCCAGATTACGAGATCTGAGAGAGCTGAACTGAGGCTGTGAGAGCGGAGTCT 780
DB 721 GGGGCTCCAGATTACGAGATCTGAGAGAGCTGAACTGAGGCTGTGAGAGCGGAGTCT 780
QY 781 GTCCTGGAACAGGCTTGTGCTGGAGCGGCTGAGCTGGGCACTGGAGTGGCTGTGGGCT 840
DB 781 GTCCTGGAACAGGCTTGTGCTGGAGCGGCTGAGCTGGGCACTGGAGTGGCTGTGGGCT 840

```

```

QY 841 CTTACATGCGCTCTGCTGCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGT 900
DB 841 CTTACATGCGCTCTGCTGCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGT 900
QY 901 CTTATTTATCACTTTGGGGGTTGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
DB 901 CTTATTTATCACTTTGGGGGTTGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
QY 961 CTTGAGATGACCTGCGCCCTGCGCCAGCCCTTACTGTGTGTGTGTGTGTGTGTGTGTGT 1020
DB 961 CTTGAGATGACCTGCGCCCTGCGCCAGCCCTTACTGTGTGTGTGTGTGTGTGTGTGTGT 1020
QY 1021 GTGCTGTGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 1060
DB 1021 GTGCTGTGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 1060

```

RESULT 2

HSNJ3280
 LOCUS Homo sapiens mRNA for 36 kDa phosphotyrosine protein.
 DEFINITION A0223280
 ACCESSION A0223280.1 GI:2780222
 VERSION phosphotyrosine protein.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Weber, J.R., Orstavik, S., Torgersen, K.M., Danbolt, N.C., Berg, S.F., Ryan, J.C., Tasken, K., Imboden, J.B. and Vaae, J.T.
 Molecular cloning of the cDNA encoding pp36, a tyrosine-phosphorylated adaptor protein selectively expressed by T cells and natural killer cells
 J. Exp. Med. 187 (7), 1157-1161 (1998)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNML

FEATURES

source

gene

CDS

```

1..1616
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /tissue_type="thymus"
  /gene="pp36"
  /codon_start=1
  /product="36 kDa phosphotyrosine protein"
  /protein_id="CA11218.1"
  /db_xref="GI:2780223"
  /db_xref="GOA:O43561"
  /db_xref="Swiss-Prot:O43561"
  /translation="MEBALTVCVLGILLPIILMLMALCVHCHRLPGSYDSTSSDSL
YPRGIQKRPHTVAWPWPAPVPTVSYPPLSQPDLPIPRSPQLGSHRTPSSRSD
GANSVASYENBEPACEDADEDDYHNHGYLVLPDSTPATSTPAASAPALSTPGIRD
SAFSMESIDYVNVVPESGESABASLDGSRVNVVSQELHGAATKTEPALSSQAEV
EEBAPPYENLQELN"

```

ORIGIN

Query Match 96.9%; Score 1027.4; DB 9; Length 1616;
 Best Local Similarity 99.9%; Pred. No. 1.4e-248;
 Matches 1026; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 GACTCTGCTTGAAGGGGCTTAAGGGGTGACGACGCTGCTCCAGCTCCCTGCAATG 60
DB 266 GACTCTGCTTGAAGGGGCTTAAGGGGTGACGACGCTGCTCCAGCTCCCTGCAATG 325

```

RESULT 3				
BC011563				
LOCUS	1472 bp	mRNA	linear	PRI 04-OCT-2003
BC011563				

```

/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="MGC:20145 IMAG:464259"
/clisue_type="Colon, adenocarcinoma"
/clone_idb="NIH MGC_15"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"

```

```

gene
1.1472
/gene="LAT"
/db_xref="LocusID:27040"
/db_xref="MIM:602354"
145_846
/gene="LAT"
/codon_start=1
/product="LAT protein"
/protein_id="AAH1563.1"
/db_xref="GI:15079451"
/db_xref="LocusID:27040"
/db_xref="MIM:602354"
/translation="MBEALIVPCVLGILLPLIAMLALCVHRLPGSYDSTSSDSL
YPRGIQKRPHPTVAWPMPAYPPTVSYPPLSQPDLPIPRSPQPGSGHRPSSRSD
GANSVASYENEPACEDADEDDYHNGPDLVLPDSTPATSTAPASAPALSTPIRD
SAFMSISIDYVNVPESEBSAASLSDSREYVNVSOELHPEAATBPAALISQBAEEV
EEBAPRYENIDELN"

ORIGIN
Query Match 95.6%; Score 1013.4; DB 9; Length 1472;
Best Local Similarity 99.9%; Pred. No. 4.8e-245;
Matches 1014; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 15 GGGGCTTAGGGGTGAGCCAGCCTGCTCCGAGCTCCCTCGAATGAGAGGCCATCT 74
DB 102 GGGGCTTAGGGGTGAGCCAGCCTGCTCCGAGCTCCCTCGAATGAGAGGCCATCT 161
QY 75 GGTCCCGTGGGTGCTGGGGCTCTGCTGCTGCGCATCTGGCATGTTGATGGGACTGTG 134
DB 162 GGTCCCGTGGGTGCTGGGGCTCTGCTGCTGCGCATCTGGCATGTTGATGGGACTGTG 221
QY 135 TGTGCACTGCCAGACTGCGAGGCTCTTACAGACAGACATCTCAGATAGTTGTATCC 194
DB 222 TGTGCACTGCCAGACTGCGAGGCTCTTACAGACAGACATCTCAGATAGTTGTATCC 281
QY 195 AAGGGGCAATCCAGTTCAAAAGGCTCTCAGAGGTCCTCCCTGGGCACTGCTTACCAACC 254
DB 282 AAGGGGCAATCCAGTTCAAAAGGCTCTCAGAGGTCCTCCCTGGGCACTGCTTACCAACC 341
QY 255 TGTACACCTCTACACCAACCCCTGAGCGAGCAGACCTGCTCCCATCCCAAGATCCCGGCA 314
DB 342 TGTACACCTCTACACCAACCCCTGAGCGAGCAGACCTGCTCCCATCCCAAGATCCCGGCA 401
QY 315 GCCCCTTGGGGGCTCCACCGGAGCGCATCTTCCCGGCGGAGTTCTGATGTTGCCAAG 374
DB 402 GCCCCTTGGGGGCTCCACCGGAGCGCATCTTCCCGGCGGAGTTCTGATGTTGCCAAG 461
QY 375 TGTGGCGAGCTACGAGAACGAGAACCGCTGTGAGGATGCAATGAGATGAGACGA 434
DB 462 TGTGGCGAGCTACGAGAACGAGAACCGCTGTGAGGATGCAATGAGATGAGACGA 521
QY 435 CTATACAAACCCAGGCTACCTGTGTGTCCTTCTGACAGCAACCCCGGCACTACGACTGC 494
DB 522 CTATACAAACCCAGGCTACCTGTGTGTCCTTCTGACAGCAACCCCGGCACTACGACTGC 581
QY 495 TGCCCCATCAAGCTCTGCACTCAGACCCCTGGCATCCGAGACAGTGCCTTCTCCATGGA 554
DB 582 TGCCCCATCAAGCTCTGCACTCAGACCCCTGGCATCCGAGACAGTGCCTTCTCCATGGA 641
QY 555 GTCCATTGATGATTAAGTAAAGTTCCGAGAGAGCGGAGAGCGCAGAGGCTCTCTGGA 614
DB 642 GTCCATTGATGATTAAGTAAAGTTCCGAGAGAGCGGAGAGCGCAGAGGCTCTCTGGA 701
QY 615 TGGCAGCCGGAGATGATGATGTCCTCCAGAACTGATCTGAGAGCGCTAGAGACTGA 674
DB 702 TGGCAGCCGGAGATGATGATGTCCTCCAGAACTGATCTGAGAGCGCTAGAGACTGA 761
QY 675 GCTTCGCGCCTGATTCAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 734
DB 762 GCTTCGCGCCTGATTCAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 821
QY 735 CGAAGATTCGAGAGAGCTGAAGCTGAGAGGCTGTGAGAGGCGAGATCTGTCTGGAACCAAG 794

```

```

DB 822 CGAAGATTCGAGAGAGCTGAAGTGAAGGCGCTGTGAGAGCGCGAGTCTGTCTGGAACCAAG 881
QY 795 CTTCGCTGGAGCGGCTGAGGCTGGGCAAGCTGGAAGTGCTGTGGGTCCTTCACATGGCGTC 854
DB 882 CTTCGCTGGAGCGGCTGAGGCTGGGCAAGCTGGAAGTGCTGTGGGTCCTTCACATGGCGTC 941
QY 855 CTGCGCTTGTCTCAGAGCTGACAAACAGCTGAGAAATCCCCGGTAATTATTAATCACTTT 914
DB 942 CTGCGCTTGTCTCAGAGCTGACAAACAGCTGAGAAATCCCCGGTAATTATTAATCACTTT 1001
QY 915 GGGGTTGGGCTGTGTGTCCCGGAAAGCTGACCTTTCGAGCGAGCTGAGAAATGACCT 974
DB 1002 GGGGTTGGGCTGTGTGTCCCGGAAAGCTGACCTTTCGAGCGAGCTGAGAAATGACCT 1061
QY 975 GCCCTGGCCCGACCGCTACTCTGTGTAAATGAATAAAGGCTCGTGTCTGTG 1029
DB 1062 GCCCTGGCCCGACCGCTACTCTGTGTAAATGAATAAAGGCTCGTGTCTGTG 1116

RESULT 4
BD107841 1224 bp DNA linear PAT 18-SEP-2002
LOCUS 36 human secreted proteins.
DEFINITION BD107841
ACCESSION BD107841.1 GI:23202659
VERSION BD107841.1 GI:23202659
KEYWORDS JP 2002500035-A/12.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1224)
RUBEN,S.M., SOPPEC,D.R., EBNER,R., LAFIEUR,D.W., NI,J.,
BREWER,L.A., OLSEN,H.S., DUAN,R.D. and ROSEN,C.A.
36 human secreted proteins
Patent: JP 2002500035-A 12 08-JAN-2002;
HUMAN GENOME SCIENCES INC
OS Homo sapiens (human)
PN JP 2002500035-A/12
PD 08-JAN-2002
PF 06-JAN-1999 JP 2000527554
PR 07-JAN-1998 US 60/070657, 07-JAN-1998 US 60/070658 PR
07-JAN-1998 US 60/070692, 07-JAN-1998 US 60/070704 PI STEVEN
W RUBEN, DANIEL R SOPPEC, REINHARD EBNER, DAVID W LAFIEUR, PI JIAN
NI,
LAURIE A BREWER, HENRIK S OLSEN, ROSANNE D DUAN, CRAIG A ROSEN PC
C12N15/09, A61K31/711, A61K38/00, A61K39/395, A61K39/395, A61K48/00,
PC A61P5/00,
A61P35/02,
A61P37/00, A61P11/06, A61P19/02, A61P29/00, A61P31/18, A61P35/00, PC
C12N15/00,
PC A61K37/02, C12N5/00
CC n equals a,t,g, or c
FH Key Location/Qualifiers
FT source 1.1224
Location/Qualifiers
1.1224
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 89.5%; Score 949.2; DB 6; Length 1224;
Best Local Similarity 98.9%; Pred. No. 8e-229;
Matches 1039; Conservative 0; Mismatches 4; Indels 8; Gaps 8;

QY 1 GACTCTGCCCTTAGAGGCGCTAGAGGCTGACAGCTGCTCCAGCTCCCTTCAGATG 60
DB 181 GACTCTGCCCTTAGAGGCGCTAGAGGCTGACAGCTGCTCCAGAGCTCCCTTCAGATG 240
QY 61 GAGGAGGCATCTGTGCTCCCTGCTGCTGGGCTCTGCTGCTGCCATCTGGGCAATG 120

```


[illegible]

VERSION	AF036506.1	GI:2828025
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 1460)	
AUTHORS	Zhang, W., Sloan-Lancaster, J., Kitchen, J., Tribble, R.P. and Samelson, L.E.	
TITLE	LAT: the ZAP-70 tyrosine kinase substrate that links T cell receptor to cellular activation	
JOURNAL	Cell 92 (1), 83-92 (1998)	
MEDLINE	98149346	
PUBMED	9489702	
REFERENCE	2 (bases 1 to 1460)	
AUTHORS	Zhang, W., Sloan-Lancaster, J., Kitchen, J., Tribble, R.P. and Samelson, L.E.	
TITLE	Direct Submission	
JOURNAL	Submitted (05-DEC-1997) Cell Biology and Metabolism Branch, National Institute of Child Health and Development, National Institute of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA	
COMMENT	LAT is a highly tyrosine phosphorylated protein, previously described as p36-38, and it associates with many signaling molecules, such as Grb2, PLC-gamma1, PI-3 kinase, cbl, Vav, and Shp-76, either directly or indirectly upon T cell activation. It is a potential type III transmembrane protein.	
FEATURES	<p>location/Qualifiers</p> <p>1..1460</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:9606"</p> <p>/cell_line="Jurkat T cells"</p> <p>1..1460</p> <p>/gene="LAT"</p> <p>/note="Linker for activation of T cell"</p> <p>79..867</p> <p>/gene="LAT"</p> <p>/note="tyrosine kinase substrate; This a alternatively spliced form of LAT"</p> <p>/codon_start=1</p> <p>/product="LAT"</p> <p>/protein_id="AAC39617.1"</p> <p>/db_xref="GI:2828026"</p> <p>/translation="MEBALIVPCVLGILLPIAMLMALCVHGRILPGSYDSTSDSLYPGQIKPQHTVAPWMPAPVPTVSYPLSPQDLPPIRSPQPLGSGHRTSPSRSDGANSVLSYENEGASGIRGAAGWGWSPSTRLTPVSLPEPACEDADEDEDYHNHG YLVVLPDSTPATSTAPASAPALSTPGRIDSAFESMEDIIDYVNPBESGSAABLDGSR EYVNVSEQLHPGAKTPEPALSSQDEAEVEEBAPEYENLELN"</p>	
ORIGIN	<p>Query Match 87.9%; Score 932; DB 9; Length 1460;</p> <p>Best Local Similarity 92.2%; Pred. No. 1.8e-224;</p> <p>Matches 1029; Conservative 0; Mismatches 0; Indels 87; Gaps 1;</p>	
QY	1	GACTCTCCCTTGAGGGGCTAGGGGTGCGAGCCAGCTGCTCGAGCTCCCTGCAAGT 60
DB	22	GACTCTCCCTTGAGGGGCTAGGGGTGCGAGCCAGCTGCTCGAGCTCCCTGCAAGT 81
QY	61	GAGGAGGCATCTGTGTCCTCCCTGTGGGTCTCTGCTGCTGCCATCTGGCCATG 120
DB	82	GAGGAGGCATCTGTGTCCTCCCTGTGGGTCTCTGCTGCTGCCATCTGGCCATG 141
QY	121	TTGATGCACTGTGTGTGCACTGCGACAGACTGCGAGGCTCTTACGACAGACATCTCA 180
DB	142	TTGATGCACTGTGTGTGCACTGCGACAGACTGCGAGGCTCTTACGACAGACATCTCA 201
QY	181	GATAGTTTGTATCCAAAGGCGATCCAGTTCAAAAGGCTCACAAGGTTGCCCCCTGGCCA 240
DB	202	GATAGTTTGTATCCAAAGGCGATCCAGTTCAAAAGGCTCACAAGGTTGCCCCCTGGCCA 261
QY	241	CTTGCTTACCACTGTGCACTCTCTTACCAACCCCTTGAAGCCAGCCAGACTGTCTCCATC 300
DB	262	CTTGCTTACCACTGTGCACTCTCTTACCAACCCCTTGAAGCCAGCCAGACTGTCTCCATC 321

QY	301	CCAAGATCCCCGACACCCCTTGGGGGGCTCCCAACGGAGCCATCTTCCCGGGGATCTT	360
Dp	322	CCAGATCCCCGGACCCCTTGGGGGGCTCCCAACGGAGCCATCTTCCCGGGGATCTT	381
QY	361	GATGTGCGCAACAGTGTGGCGAGCTACGAGAACGAG-----	396
Dp	382	GATGTGCGCAACAGTGTGGCGAGCTACGAGAACGAGGGTGGCTGGGATCCGAGTGCC	441
QY	397	-----	396
Dp	442	CAGGCTGGGTGGGAGTCTGGGGTCCGCTTGAGACTAGGCTGACCCCTGTGTCTTACC	501
QY	397	---GAAACAGCCTGTGAGAGATGCAGATGAGATGAGACGACTATCACAAACCGAGCTAC	453
Dp	502	CCAGAACCGACTGTGAGAGATGCAGATGAGATGAGACGACTATCACAAACCGAGCTAC	561
QY	454	CTGTGTGTCTTCTTGACAGCACCCCGGACCTAGCACTGTCTGCCCATCAGCTCTTGCA	513
Dp	562	CTGTGTGTCTTCTTGACAGCACCCCGGACCTAGCACTGTCTGCCCATCAGCTCTTGCA	621
QY	514	CTCAGCAACCCCTGGATCCGAGACAGTGCCTTCTCCATGAGATCCATGATGATTAAGTGC	573
Dp	622	CTCAGCAACCCCTGGATCCGAGACAGTGCCTTCTCCATGAGATCCATGATGATTAAGTGC	681
QY	574	AACGTTCCGAGAGCGGGGAGAGCCCAAAAGCGCTCTGATATGACAGCGGGAGATGTG	633
Dp	682	AACGTTCCGAGAGCGGGGAGAGCGCCAAAGCGCTCTGATATGACAGCGGGAGATGTG	741
QY	634	AATGTGTCCAGGAATCTGCATCTTGAGAGCGCTTAAGACTGAGCCTGCCCTTGAGTTCC	693
Dp	742	AATGTGTCCAGGAATCTGCATCTTGAGAGCGCTTAAGACTGAGCCTGCCCTTGAGTTCC	801
QY	694	CAGAGAGCAGAGGAAGTGAAGGAAGAGGGGGCTCCAGATTTCAGAAATCTGCAGAGCTG	753
Dp	802	CAGAGAGCAGAGGAAGTGAAGGAAGAGGGGGCTCCAGATTTCAGAAATCTGCAGAGCTG	861
QY	754	AACCTAGGGGCTGTGAGAGCGGAGTCTGTCTTGGAACCAAGGCTTGCTGGAGCGGCTAG	813
Dp	862	AACCTAGGGGCTGTGAGAGCGGAGTCTGTCTTGGAACCAAGGCTTGCTGGAGCGGCTAG	921
QY	814	CTGGGCACTGGAAGTGGCTCTGGGGGTCTCACATGAGGGTCTTGCCCTTGCTCCAGCTGT	873
Dp	922	CTGGGCACTGGAAGTGGCTCTGGGGGTCTCACATGAGGGTCTTGCCCTTGCTCCAGCTGT	981
QY	874	ACAACAGCCTGAGAAATCCCCCGGTAATCTTATCACTTTGGGGGTTGGGCTGTGTCC	933
Dp	982	ACAACAGCCTGAGAAATCCCCCGGTAATCTTATCACTTTGGGGGTTGGGCTGTGTCC	1041
QY	934	CCGAAAGCTGTGACCTTCTGACGAGCCTTGAAATGACCTTGCCCTTGAGCCCAAGCCTAC	993
Dp	1042	CCGAAAGCTGTGACCTTCTGACGAGCCTTGAAATGACCTTGCCCTTGAGCCCAAGCCTAC	1101
QY	994	TCTGTGTAAATGAATAAAGGCTGTGGTGTGTCTGTG 1029	
Dp	1102	TCTGTGTAAATGAATAAAGGCTGTGGTGTGTGTG 1137	

Ieshi, Y., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahori, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuna, M., Shiraori, A., Sudo, H., Hosokita, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Kasuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Niinomiya, K., Ishibashi, T., Yamaishi, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yoshida, M., Hotta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiyasu, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Muraishi, K., Yuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terasahima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kamigaki, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikil, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsunura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togaishi, T., Oyama, M., Hara, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuno, Y., Yamaishi, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T., and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs
 Nat. Genet. 36 (1), 40-45 (2004)

14702039

JOURNAL
 PUBMED
 REFERENCE

AUTHORS

2

Ieshibashi, T., Kanehori, K., Yoshida, M., Watanabe, S., Ishida, S., Ono, Y., Hotta, T., Hiraoka, S., Murakawa, K., Takiguchi, S., Kusano, J., Watanabe, M., Fujimori, K., Tanai, A., Ichida, M., Yamaishi, H., Chiba, Y., Sugiyama, T., Ito, R., Otsubuki, T., Sato, H., Ota, T., Wakamatsu, A., Ieshi, S., Yamamoto, Y., Isono, Y., Kawai, Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Matsumoto, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuna, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Sugano, S., Nagahori, K., Masuno, Y., Nagai, K., and Isogai, T. NEO human cDNA sequencing project
 Unpublished
 3 (bases 1 to 2443)
 Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isogai, FUJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 297-0812, Japan
 (E-mail:genominfo@hri.co.jp, Tel:81-338-52-3975, Fax:81-438-52-3986)
 NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'-4'-3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

COMMENT

RESULT 6	AK093815	LOCUS	AK093815	2443 bp	mRNA	linear	PRI 30-JAN-2004
DEFINITION	Homo sapiens CDNA FLJ36496 fig. clone THYMZ018819.	ACCESSION	AK093815	AK093815.1	GI:21752752	oligo capping: figs (full insert sequence).	
VERSION	AK093815.1	KEYWORDS	oligo capping: figs (full insert sequence).				
SOURCE	Homo sapiens (human)	ORGANISM	Homo sapiens				
REFERENCE	1	Authors	Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Oobayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="THYM2018819"
/issue_type="chymusa"
/clone_lib="THYM2"
/notes="Cloning vector: pME18SFL3"

ORIGIN

Query Match      45.2%; Score 479.6; DB 9; Length 2443;
Best Local Similarity 99.2%; Pred. No. 4.e-110;
Matches 482; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      544  TTCTCCATGGAGTCCATTGATGATTAGCTGAACCTTCGGAGAGCGGAGACGGCAGAA 603
      |||

```

```

Db      1955  TCACAGTGGAGTCCATTGATGATTACGTGAACGTTCGAGAGCGGGAGAGCCAGAA  2014
Qy      604  GGTCTCTGGATGAGAGCCGGAGATATGTAATGTGTCCAGAACTGATCTGTGAGAG  663
Db      2015  GGTCTCTGGATGAGAGCCGGAGATATGTAATGTGTCCAGAACTGATCTGTGAGAG  2074
Qy      664  GCTAAGACTGAGACCTGCGCCCTGTGATTCCAGAGAGCAGAGAAATGAGAGAGGG  723
Db      2075  GCTAAGACTGAGACCTGCGCCCTGTGATTCCAGAGAGCAGAGAAATGAGAGAGGG  2134
Qy      724  GCTCAAGATTACGAAATCTGAGAGAGCTGAATGAGAGGCTGTGAGAGCCGAGTCTGTC  783
Db      2135  GCTCAAGATTACGAAATCTGAGAGAGCTGAATGAGAGGCTGTGAGAGCCGAGTCTGTC  2194
Qy      784  CTGGAACCAAGGCTGCTGCGGAGAGCTGAGCTGAGCTGAGAGTGGCTCTGGGGTCTCT  843
Db      2195  CTGGAACCAAGGCTGCTGCGGAGAGCTGAGCTGAGCTGAGAGTGGCTCTGGGGTCTCT  2254
Qy      844  CACATGCGCTCTGCGCTGCTGCTCAGAGCTGAGCAACAGAGCTGAGAAATCCCGCTAATT  903
Db      2255  CACATGCGCTCTGCGCTGCTGCTCAGAGCTGAGCAACAGAGCTGAGAAATCCCGCTAATT  2314
Qy      904  ATTATCACTTTGGGGTGTGCGCTGTGTGTCGCCGAAAGCTCTGACCTTCTGACGAGCCT  963
Db      2315  ATTATCACTTTGGGGTGTGCGCTGTGTGTCGCCGAAAGCTCTGACCTTCTGACGAGCCT  2374
Qy      964  GAGAAATGACCTGCTGCGCCCTGAGCCCAAGCCCTACTCTGTGTATTAAGAAAGGCTGTG  1023
Db      2375  GAGAAATGACCTGCTGCGCCCTGAGCCCAAGCCCTACTCTGTGTATTAAGAAAGGCTGTG  2434
Qy      1024  TCTGTG 1029
Db      2435  TCTGTG 2440

RESULT 7
LOCUS   AR380020 528 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 565 from patent US 6607879.
ACCESSION AR380020
VERSION  AR380020.1 GI:40087654
KEYWORDS
SOURCE  Unknown.
ORGANISM  Unclassified.
REFERENCE  1 (bases 1 to 528)
AUTHORS  Cocks,B.G., Stuart,S.G. and Selthamer,J.J.
TITLE    Compositions for the detection of blood cell and immunological
          response gene expression
          Patent: US 6607879-A 565 19-AUG-2003;
          Location/Qualifiers
            1..528
              /organism="unknown"
              /mol_type="genomic DNA"

ORIGIN
Query Match 39.3%; Score 416.8; DB 6; Length 528;
Best Local Similarity 99.3%; Pred. No. 2,8e-94;
Matches 429; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

```

Db      181  CTGTCTGGAACCAAGCTTGCCTGGGAACGGGTGAGCTGGGAGCTGGAAGTGGCTGG  240
Qy      838  GGTCTCTCAATGAGGCTCTGCGCTTGTCTCAGAGCTGAGCAACAGAGCTGAGAAATCCCGCG  897
Db      241  GGTCTCTCAATGAGGCTCTGCGCTTGTCTCAGAGCTGAGCAACAGAGCTGAGAAATCCCGCG  300
Qy      898  TAACTTATATACATTGGGGGTTGGGCTGTGTCCCGGAAGGCTCTGACCTTGTGAGC  957
Db      301  TAACTTATATACATTGGGGGTTGGGCTGTGTCCCGGAAGGCTCTGACCTTGTGAGC  360
Qy      958  CAGCTGAGAAATGAGAGCTGCGCCCTGAGCCCAAGCCCTACTCTGTGTATTAAGAAATGAGCCCTG  1017
Db      361  CAGCTGAGAAATGAGAGCTGCGCCCTGAGCCCAAGCCCTACTCTGTGTATTAAGAAATGAGCCCTG  420
Qy      1018  CGTGTGTCTGTG 1029
Db      421  CGTGTGTCTGTG 432

RESULT 8
LOCUS   BC052340 1260 bp mRNA linear ROD 30-JUN-2004
DEFINITION Mus musculus linker for activation of T cells, mRNA (cDNA clone
ACCESSION BC052340
VERSION  BC052340.1 GI:30354172
KEYWORDS
SOURCE  MGC.
ORGANISM  Mus musculus (house mouse)
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1260)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,W.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Uedl,T.B., Toshilyuk,I.S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaby,S.U., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smalins,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
1247932
2 (bases 1 to 1260)
Strausberg,R.
Direct Submission
Submitted (02-MAY-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cga@pds-remail.nih.gov
Tissue Procurement: Drs. Dennis Taub, Dan Longo (NIA, USA),
Jonathan Keller (NCI, USA)
cDNA Library Preparation: Yulan Piao and Minoru Ko (National
Institute on Aging, NIH: http://jgsum.grc.nia.nih.gov/cDNA/)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMIL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;

```


AUTHORS Zhang, W., Sloan-Lancaster, J., Kitchen, J., Trible, R. P. and Samelson, L. E.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-1997) Cell Biology and Metabolism Branch, National Institute of Child Health and Development, National Institute of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA
COMMENT LAT is a highly tyrosine phosphorylated protein, previously described as p36-38, and it associates with many signaling molecules, such as Grb2, Plc-gamma1, PI-3 kinase, cbl, Vav, and SLP-76, either directly or indirectly upon T cell activation. It is a potential type III transmembrane protein.
FEATURES
 source 1..1260
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /tissue_type="thymus"
 /dev_stage="adult"
 1..1260
 /gene="LAT"
 /note="linker for activation of T cells"
 292..1020
 /gene="LAT"
 /note="tyrosine kinase substrate"
 /codon_start=1
 /product="LAT"
 /protein_id="AAC40054.1"
 /db_xref="GI:2828028"
 /translation="MEADALSPVGLGLLPLVLTLLALCYRCRELPVSYDSTSTESLYPRSLIPQITVPRTPAVSPLVPSPPRLPLIPSPQPLGSHMPSQNSDDANSVASVENOSPACKNVDADEDEDYPNGYLVLPDSPPAAYVVSAPVPSNPDLGDAFVSCEDEYVNPESBESAASLDSREYVNSPEQQTPTRALIASVNDQVEBGESEBGEVDSBEPDYENLQELN"

ORIGIN

Query Match 37.8%; Score 400.4; DB 10; Length 1260;
 Best Local Similarity 67.6%; Pred. No. 4.1e-90;
 Matches 698; Conservative 0; Mismatches 291; Indels 43; Gaps 8;

QY 26 GTGCAGCAGCAGCTGCTCCGAGCTCCCTCGAATGAGAGGCGCATCTGTCCTCGCG 85
 DB 260 GTGCAGCAGGCGCTGTGCGAGGTCTCTGAGATGGAACAGCGCTTGAGCCCGGTG 319
 QY 86 TGCTGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 142
 DB 320 GGTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379
 QY 143 GCCACAGACTGCGAGGCTCTGAGCAGACATCTCAGATAGTTGTATCCAGAGGGGA 202
 DB 380 GCGGAGGTGCGAGTCTCTGAGCAGACATCTCAGAGATTTGTATCCAGAGAGCA 439
 QY 203 TCCAGTTCAACGGGCTCAACAGGTGCCCCCTGGCCACTG-----CTTACCCACTG 256
 DB 440 TCCATATCAAGCAGCTCAATTAACCGTCCCCGAGACACTGCTGTTCTTACCTCTAG 499
 QY 257 TTAACCTCTTACCCAGGCTGAGCAGCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 316
 DB 500 TCACTTCTTCCACCCCTGAGGACAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 559
 QY 317 CCCTTGGGGGCTCCCAACGAGCGCATCTTCCCGGGGGATCTGATGAGCCACAGTG 376
 DB 560 CCCTTGGGGGCTCCCAACGAGCGCATCTTCCCGAGAAATTCAGATGATGCCAAGAGTG 619
 QY 377 TGGCGAGCTAGAGAAAGAGAAACAGCTGTGAGAGATCAGATGAGATGAGAGAGACT 436
 DB 620 TGGCAAGCTAGAGAAACAGAGCCAGCTGTAAAGATGATGAGATGAGATGAGAGAG 679
 QY 437 ATCAACAACC---AGCTACTGTGTGTGTCTTCTTACAGCAGCAGCCGGGCACTAGACTG 493
 DB 680 AGGACTATCCAAAGGCTACTAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 739
 QY 494 CTGCCCATCAGCTCTGCACTCAGACCCCTGAGCATCCGAGACAGTGCCTTCTCCATGG 553

DB 740 TTGTCTCTCTGCTCTGCTGCTGCTAGAACCTGACTTGGAGACAGTGCCTTCTGTGG 799
 QY 554 AGTCATTTGATATACGTAACGTAACGTTCCGGAAGCGGGAGCGGAGAGGCTCTCTGG 613
 DB 800 AGTCGTGTAATTAACGTAATGTTCTCGAAGTGAAGAGAGCGAGAGGCGTCTCTGG 859
 QY 614 ATGGACGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 673
 DB 860 ATGGACGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 916
 QY 674 AGCTGCGCGCTGATGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 733
 DB 917 AGCTGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 976
 QY 734 AGGAGATGAG 793
 DB 977 GAG 1036
 QY 794 GCTTGGCGGAG 853
 DB 1037 TCTGTCTG-----AACTGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1088
 QY 854 CCTGCGCTTGTCTCAGCTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 913
 DB 1089 CTCTGCTTGCACAGCTGA-----GAATCTTCCCTTAATTATGTCACCTT 1136
 QY 914 TGGGTTGCGGCTGT 973
 DB 1137 TGGGTTGCGGCTGT 1194
 QY 974 TCCCTGCGGCGAGCGCTACTGCTGATATGATATGATATGATATGATATGATATGATATG 1033
 DB 1195 CTGTTTCCAGCAGAC-----ATGTCATGAGATTAAGCCATGTACATTAATAAAAAA 1248
 QY 1034 AAAAAAAAAAAAAA 1045
 DB 1249 AAAAAAAAAAAAAA 1260

RESULT 10
 LOCUS RNJ001184 1188 bp mRNA linear ROD 10-FEB-1999
 DEFINITION Rattus norvegicus mRNA for 36 kDa phosphoprotein.
 ACCESSION AJ001184
 VERSION AJ001184.1 GI:3021582
 KEYWORDS 36 kDa phosphoprotein.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1
 REFERENCE
 AUTHORS Weber, J.R., Orstavik, S., Torgersen, K.M., Danbolt, N.C., Berg, S.F., Ryan, J.C., Taeken, K., Imboden, J.B. and Vaae, J.T.
 TITLE Molecular cloning of the cDNA encoding pp36, a tyrosine-phosphorylated adaptor protein selectively expressed by T cells and natural killer cells
 JOURNAL J. Exp. Med. 187 (7), 1157-1161 (1998)
 MEDLINE 98197173
 PUBMED 9529333
 REFERENCE 2 (bases 1 to 1188)
 AUTHORS Orstavik, S.
 TITLE Direct Submission
 JOURNAL Submitted (08-JAN-1998) Orstavik S., Institute of medical biochemistry, University of Oslo, pb 1112, Blindern, N-0317 Oslo, NORWAY
 FEATURES
 source 1..1188
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="PVG"
 /db_xref="taxon:10116"
 /clone="151"

Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

FEATURES

source
Location/Qualifiers
1. .131839
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="CTB-134H23"

ORIGIN

Query Match 25.0%; Score 265; DB 9; Length 131839;
Best Local Similarity 98.2%; Pred. No. 8.9e-56;
Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 757 TGAGGGGCTGTGGAGGCGGAGTCTGTCTGGAACGAGCTTGGCTGGAGGCGCTGAGCTG 816
Db 110544 TGTGTTTCAGTGGAGGCGGAGTCTGTCTGGAACGAGCTTGGCTGGAGGCGCTGAGCTG 110485
QY 817 GGCAGCTGGAAGTGGCTGTGGGGTCTTCACATGGCGTCTCCCTTGTCTCCAGCTGACA 876
Db 110484 GGCAGCTGGAAGTGGCTGTGGGGTCTTCACATGGCGTCTCCCTTGTCTCCAGCTGACA 110425
QY 877 ACAGCTGGAAGAAATCCCCCGTAATTATTCACATTGGGGTTTCGGCTGTGTCCCGC 936
Db 110424 ACAGCTGGAAGAAATCCCCCGTAATTATTCACATTGGGGTTTCGGCTGTGTCCCGC 110365
QY 937 AACGCTTGACCTTCTGACGAGCTGAGATGACCTGCGGCCCGCCAGCCCTACTCT 996
Db 110364 AACGCTTGACCTTCTGACGAGCTGAGATGACCTGCGGCCCGCCAGCCCTACTCT 110305
QY 997 GTGTAATAGATAAAGGCTGCTGTGTCTGTG 1029
Db 110304 GTGTAATAGATAAAGGCTGCTGTGTCTGTG 110272

RESULT 13

AC109460 163424 bp DNA linear PRI 12-OCT-2002
LOCUS Homo sapiens chromosome 16 clone RP11-264B17, complete sequence.
DEFINITION
ACCESSION AC109460 GI:23915447
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 163424)
DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
TITLE
JOURNAL
AUTHORS
REFERENCE

2 (bases 1 to 163424)
DOE Joint Genome Institute.
TITLE
JOURNAL
AUTHORS
REFERENCE

3 (bases 1 to 163424)
DOE Joint Genome Institute.
TITLE
JOURNAL
AUTHORS
REFERENCE

Submitted (03-SBP-2002) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 163424)
DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.

TITLE
JOURNAL
AUTHORS
REFERENCE

Submitted (12-OCT-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Oct 12, 2002 this sequence version replaced gi:22657513.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center and Los Alamos
National Laboratory
www.sngc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.3.

FEATURES

source
Location/Qualifiers
1. .163424
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-264B17"

ORIGIN

Query Match 25.0%; Score 265; DB 9; Length 163424;
Best Local Similarity 98.2%; Pred. No. 9e-56;
Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 757 TGAGGGGCTGTGGAGGCGGAGTCTGTCTGGAACGAGCTTGGCTGGAGGCGCTGAGCTG 816
Db 122732 TGTGTTTCAGTGGAGGCGGAGTCTGTCTGGAACGAGCTTGGCTGGAGGCGCTGAGCTG 122791
QY 817 GGCAGCTGGAAGTGGCTGTGGGGTCTTCACATGGCGTCTCCCTTGTCTCCAGCTGACA 876
Db 122792 GGCAGCTGGAAGTGGCTGTGGGGTCTTCACATGGCGTCTCCCTTGTCTCCAGCTGACA 122851
QY 877 ACAGCTGGAAGAAATCCCCCGTAATTATTCACATTGGGGTTTCGGCTGTGTCCCGC 936
Db 122852 ACAGCTGGAAGAAATCCCCCGTAATTATTCACATTGGGGTTTCGGCTGTGTCCCGC 122911
QY 937 AACGCTTGACCTTCTGACGAGCTGAGATGACCTGCGGCCCGCCAGCCCTACTCT 996
Db 122912 AACGCTTGACCTTCTGACGAGCTGAGATGACCTGCGGCCCGCCAGCCCTACTCT 122971
QY 997 GTGTAATAGATAAAGGCTGCTGTGTCTGTG 1029
Db 122972 GTGTAATAGATAAAGGCTGCTGTGTCTGTG 123004

RESULT 14

AC138925 187788 bp DNA linear HTG 21-JAN-2003
LOCUS Homo sapiens chromosome 16 clone RP11-664F22, WORKING DRAFT
DEFINITION
ACCESSION AC138925
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 187788)
DOE Joint Genome Institute.
TITLE
JOURNAL
AUTHORS
REFERENCE

2 (bases 1 to 187788)
DOE Joint Genome Institute.
TITLE
JOURNAL
AUTHORS
REFERENCE

Submitted (21-JAN-2003) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI

Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 1549622
Center Clone name: RPCT-11_664F22

Summary Statistics
Consensus quality: 186934 bases at least Q40
Consensus quality: 187077 bases at least Q30
Consensus quality: 187284 bases at least Q20
Estimated insert size: 180000; agarose-fp estimation
Estimated insert size: 187688; sum-of-ctnigs estimation
Quality coverage: 11.48 in Q20 bases; agarose-fp estimation
Quality coverage: 11.01 in Q20 bases; sum-of-ctnigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 34623: contig of 34623 bp in length
* 34624 34723: gap of unknown length
* 34724 187788: contig of 153065 bp in length.

Location/Qualifiers
1. 187788
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-664F22"
/clone_lib="RPCT human BAC library 11"

ORIGIN

Query Match 25.0%; Score 265; DB 2; Length 187788;
Best Local Similarity 98.2%; Pred. No. 9.1e-56;
Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 757 TGAAGGCTGTGAGAGCCGAGTGTCTGTGGAACGAGCTTGCGAGCTG 816
DB 111946 TGTGTTTCAGTGAAGCCGAGTGTCTGTGGAACGAGCTTGCGAGCTG 111887

QY 817 GGCAGCTGGAAGTGGCTGTGGGGTCTCTCATGAGCGCTTGCCTCCAGCCGAGA 876
DB 111886 GGCAGCTGGAAGTGGCTGTGGGGTCTCTCATGAGCGCTTGCCTCCAGCCGAGA 111827

QY 877 ACAGCTGGAAGTGGCTGTGGGGTCTCTCATGAGCGCTTGCCTCCAGCCGAGA 936
DB 111826 ACAGCTGGAAGTGGCTGTGGGGTCTCTCATGAGCGCTTGCCTCCAGCCGAGA 111767

QY 937 AACGCTGTGACCTTCTGACGAGCCTGGAATGACCTGCGCCGAGCCCTACTCT 996
DB 111766 AACGCTGTGACCTTCTGACGAGCCTGGAATGACCTGCGCCGAGCCCTACTCT 111707

QY 997 GTGTATAGAAATTAAGCCTGTGCTGTCTGTG 1029
DB 111706 GTGTATAGAAATTAAGCCTGTGCTGTCTGTG 111674

RESULT 15
AC139247
DEFINITION Homo sapiens chromosome 16 clone CTD-2553D4, WORKING DRAFT
SEQUENCE, 18 unordered pieces.
AC139247
AC139247.1 GI:28008725
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 219646)

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 219646)
DOE Joint Genome Institute.
Direct Submission
Submitted (29-JAN-2003) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 810409
Center Clone name: CTD-2553D4

Summary Statistics
Consensus quality: 211019 bases at least Q40
Consensus quality: 212006 bases at least Q30
Consensus quality: 212932 bases at least Q20
Estimated insert size: 175000; agarose-fp estimation
Estimated insert size: 217946; sum-of-ctnigs estimation
Quality coverage: 31.21 in Q20 bases; agarose-fp estimation
Quality coverage: 25.06 in Q20 bases; sum-of-ctnigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1100: contig of 1100 bp in length
* 1101 1200: gap of unknown length
* 1201 2326: contig of 1126 bp in length
* 2327 2426: gap of unknown length
* 2427 3448: contig of 1022 bp in length
* 3449 3548: gap of unknown length
* 3549 4793: contig of 1245 bp in length
* 4794 4893: gap of unknown length
* 4894 5992: contig of 1099 bp in length
* 5993 6092: gap of unknown length
* 6093 8163: contig of 2071 bp in length
* 8164 8263: gap of unknown length
* 8264 9387: contig of 1124 bp in length
* 9388 9487: gap of unknown length
* 9488 10873: contig of 1386 bp in length
* 10874 10973: gap of unknown length
* 10974 12188: contig of 1215 bp in length
* 12189 12288: gap of unknown length
* 12289 13292: contig of 1004 bp in length
* 13293 13392: gap of unknown length
* 13393 14698: contig of 1306 bp in length
* 14699 14798: gap of unknown length
* 14799 16556: contig of 1758 bp in length
* 16557 16656: gap of unknown length
* 16657 18479: contig of 1823 bp in length
* 18480 18579: gap of unknown length
* 18580 22438: contig of 3859 bp in length
* 22439 22538: gap of unknown length
* 22539 40064: contig of 17526 bp in length
* 40065 40164: gap of unknown length
* 40165 53730: contig of 13566 bp in length
* 53731 53830: gap of unknown length
* 53831 89191: contig of 35361 bp in length
* 89192 89291: gap of unknown length
* 89292 219646: contig of 130355 bp in length.

Location/Qualifiers
1. 219646
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source

/chromosome="16"
/clone="CTD-2553D4"
/clone_lib="Caltech human BAC library D"

ORIGIN

Query Match 25.0%; Score 265; DB 2; Length 219646;
Best Local Similarity 98.2%; Pred. No. 9.2e-56;
Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 757 TGAGGGCTGTGAGAGCCGAGTCTGTCTGGAACAGGCTTGCTGGGAGCGCTGAGCTG 816
|||
Db 130373 TGTGTTTCAGTGAAGCCGAGTCTGTCTGGAACAGGCTTGCTGGGAGCGCTGAGCTG 130432

QY 817 GGCAGCTGGAAGTGGCTCTGGGGTCTTCACATGGCGTCTGCCCTTGCTCCAGCCTGACA 876
|||
Db 130433 GGCAGCTGGAAGTGGCTCTGGGGTCTTCACATGGCGTCTGCCCTTGCTCCAGCCTGACA 130492

QY 877 ACAGCCTGAGAAATCCCCCGTAATTATATCACTTTGGGGTTGCGCCTGTGTCCCCG 936
|||
Db 130493 ACAGCCTGAGAAATCCCCCGTAATTATATCACTTTGGGGTTGCGCCTGTGTCCCCG 130552

QY 937 AACGCTGTGACCTTCTGAGCGCAGCCTGAGATGACCTGCGCCCGCCAGCCTACTCT 996
|||
Db 130553 AACGCTGTGACCTTCTGAGCGCAGCCTGAGATGACCTGCGCCCGCCAGCCTACTCT 130612

QY 997 GTGTATAGATTAAGGCTGCGTGTGTGTG 1029
|||
Db 130613 GTGTATAGATTAAGGCTGCGTGTGTGTG 130645

Search completed: November 22, 2004, 10:12:47
Job time : 6516 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 04:26:10 ; Search time 985 Seconds
(without alignments)
6287.442 Million cell updates/sec

Title: US-09-597-920B-1
Perfect score: 1060
Sequence: 1 gactctgccttcgtgagggcc.....aaaaaaaaaaaaaaaaaa 1060

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N Geneseq_23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1048	98.9	1059	2	AAx89072 Human LAT
2	949.2	89.5	1224	2	AAZ06221 Human sec
3	932	87.9	1488	12	ADQ22892 Human sec
4	921	86.9	1461	2	AAx89073 Human LAT
5	454	42.8	518	12	ADQ18376 Human LAT
6	416.8	39.3	528	11	AD131239 Human CDN
7	400.4	37.8	1260	2	AAx89074 Human sec
8	383	36.4	512	4	AAK57405 Human imm
9	68	6.4	319	4	AAK57403 Human imm
10	54.6	5.2	6307	9	ADA19460 Mouse LAT
11	53.2	5.0	2000	8	ADA71938 Rice gene
12	49.6	4.7	2968	4	AAc78190 Human can
13	49.6	4.7	2968	4	AAH34408 Human can
14	49.6	4.7	2968	4	AAH34408 Human can
15	48.6	4.6	375	4	AAI18830 Human pol
16	48	4.5	538	10	ADG32840 Human DNA
17	47.6	4.5	463	4	AAI192452 Human DNA
18	47.4	4.5	3196	6	AAZ36812 DNA encod
19	47.4	4.5	3196	6	AAH63368 Mouse bra
20	46.8	4.4	2000	8	ADA71938 Rice gene
21	46.6	4.4	227	5	ABV56539 Human pro

22	46	4.3	203	5	ABV47873 Human pro
23	46	4.3	477	9	ACH23695 Human adu
24	45.6	4.3	468	9	ACH34618 Human end
25	44.8	4.2	173	4	AAH60588 Human can
26	44.8	4.2	189	4	AAH60389 Human can
27	44.8	4.2	189	4	AAH60389 Human can
28	44.8	4.2	1362	4	AAH34417 Human can
29	44.8	4.2	1362	4	AAH34417 Human can
30	44.6	4.2	2103	12	ADQ22264 Human trf
31	44.6	4.2	424	6	ABU78339 Human ova
32	44.4	4.2	472	9	ACH29376 Human adu
33	44.4	4.2	234	4	AAI15324 Human bre
34	44.4	4.2	1694	2	AAV33190 Secreted
35	44.2	4.2	2718	9	ACF05973 Human car
36	44.2	4.2	1338	2	AAx97959 Human sec
37	44.2	4.2	1338	8	ADA39939 Human sec
38	44.2	4.2	1338	8	ACC50518 Human sec
39	44.2	4.2	1338	10	ADCT3577 Human CDN
40	44.2	4.2	1338	10	ADA56129 Gene enco
41	44.2	4.2	1529	12	ADQ24297 Human sof
42	44.2	4.2	1537	10	AAI53565 CDNA enco
43	44.2	4.2	1677	6	ABU89839 Human pol
44	44.2	4.2	2891	5	AAH75804 DNA encod
45	44	4.2	139	4	AAI24170 Human bre

ALIGNMENTS

RESULT 1
AAx89072
ID AAX89072 standard; DNA; 1059 BP.

AAx89072;
14-SEP-1999 (first entry)

Human LAT (linker for activation of T cells) protein encoding DNA.
LAT, tyrosine kinase, linker for activation of T cell, TCR, human;
T-cell receptor; TCR signalling pathway; neoplasia; inflammation;
hypersensitivity; allergy; microbial infection; genetic disease;
autoimmune disease; graft rejection; modulator; ss.

OS Homo sapiens.

PN WO9332627-A2.

XX 01-JUL-1999.

PD 23-DEC-1998; 98WO-US027400.

PF 23-DEC-1997; 97US-0068690P.

PR 23-DEC-1997; 97US-0068690P.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Samejson LE, Zhang W;

XX WPI, 1999-418926/35.

DR P-PSDB; AAY27120.

XX Linker for activation of T cell protein used to, e.g. screen for
modulators of T cell signaling.

XX Claim 8; Fig 7A; 125bp; English.

CC The invention relates to a protein tyrosine kinase substrate LAT (linker
for activation of T cells) protein. Modulation of interaction between LAT
and the T-cell receptor (TCR) affects the TCR signalling pathway. LAT is
a substrate for tyrosine kinases and becomes phosphorylated after TCR
engagement, resulting in recruitment of other signalling molecules. LAT
is used to identify and test (ant)agonists of tyrosine kinase signalling
pathways, i.e. modulation of interaction between tyrosine kinase

CC substrates and intracellular ligands or between these ligands and other
 CC members of the pathway, including identification of downstream signalling
 CC proteins, particularly in immune system cells. These modules are
 CC potentially useful as drugs and diagnostic agents, particularly for
 CC diseases that involve undesirable cell proliferation, differentiation,
 CC growth or cell energy, e.g. neoplasia, inflammation, hypersensitivity/
 CC allergy, microbial infection, metabolic, genetic or autoimmune diseases,
 CC graft rejection. LAT is also used to generate specific antibodies, used
 CC for detection of LAT. Nucleic acid that encodes LAT, or its fragments,
 CC are used to identify homologous sequences in other species, to detect the
 CC LAT gene and as sources of antisense therapeutics. Modulators of LAT are
 CC potentially more specific and less toxic than known immunosuppressants
 CC such as cyclosporin. The present sequence represents the nucleotide
 CC sequence of human LAT
 CC
 CC

Sequence 1059 BP; 227 A; 330 C; 301 G; 201 T; 0 U; 0 Other;

Query Match 98.9%; Score 1048; DB 2; Length 1059;
 Best Local Similarity 99.9%; Pred. No. 2,2e-233;
 Matches 1059; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GACTCTGCTTGAAGGGGCTTGAAGGCTGACAGCCCTGCTCCAGCTCCCTGCAAGT 60
 DB 1 GACTCTGCTTGAAGGGGCTTGAAGGCTGACAGCCCTGCTCCAGCTCCCTGCAAGT 60
 QY 61 GAGGAGGCACTCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 DB 61 GAGGAGGCACTCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 119
 QY 121 TTGATGCACTGTGTGCACTGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 DB 120 TTGATGCACTGTGTGCACTGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 179
 QY 181 GATGCTTGTATCCAAAGGGGATCAAGTTCAAAGGCTTCAAGGCTTCAAGGCTTCAAGG 240
 DB 180 GATGCTTGTATCCAAAGGGGATCAAGTTCAAAGGCTTCAAGGCTTCAAGGCTTCAAGG 239
 QY 241 CTTGCTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 DB 240 CTTGCTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 299
 QY 301 CCAAGATCCCGGAGCCCTTGGGGGCTCCCAAGGAGCCGATTTCCCGGGGATTTCT 360
 DB 300 CCAAGATCCCGGAGCCCTTGGGGGCTCCCAAGGAGCCGATTTCCCGGGGATTTCT 359
 QY 361 GATGCTGCAACATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 DB 360 GATGCTGCAACATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
 QY 421 GAGGATGAGACGACTATCAACAACCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 DB 420 GAGGATGAGACGACTATCAACAACCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
 QY 481 GCCACTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 DB 480 GCCACTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 QY 541 GCTTCTCCATGAGTTCATTGATGATTAAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 600
 DB 540 GCTTCTCCATGAGTTCATTGATGATTAAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 599
 QY 601 GAAGCGTCTGATGAGGAGCGGGAGTATGATGATGATGATGATGATGATGATGATGAT 660
 DB 600 GAAGCGTCTGATGAGGAGCGGGAGTATGATGATGATGATGATGATGATGATGATGAT 659
 QY 661 GCGGCTTAAGTGAAGCTGCGGCTTGAAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAG 720
 DB 660 GCGGCTTAAGTGAAGCTGCGGCTTGAAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAG 719
 QY 721 GGGGCTTCAATTAAGGAGTTCAGGAGTGAAGTGAAGGAGGAGGAGGAGGAGGAGGAG 780
 DB 720 GGGGCTTCAATTAAGGAGTTCAGGAGTGAAGTGAAGGAGGAGGAGGAGGAGGAGGAG 779

QY 781 GTCTGGAACACAGCTTGGAGCGGCTGAGCTGAGCTGAGGAGTGGCTTCTGGGGT 840
 DB 780 GTCTGGAACACAGCTTGGAGCGGCTGAGCTGAGCTGAGGAGTGGCTTCTGGGGT 839
 QY 841 CTTCACTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 DB 840 CTTCACTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899
 QY 901 CTTATTAATCACTTGGGGGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 DB 900 CTTATTAATCACTTGGGGGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
 QY 961 CTTGAGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 DB 960 CTTGAGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019
 QY 1021 GTGCTGTGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 1060
 DB 1020 GTGCTGTGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 1059

RESULT 2

AAZ06221 standard; DNA; 1224 BP.
 AAZ06221;
 30-SEP-1999 (first entry)
 Human secreted protein gene No. 3.
 Human; secreted protein; fusion protein; gene therapy; protein therapy;
 diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 developmental abnormality; fetal deficiency; blood; allergy; renal; ds;
 immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 Homo sapiens.
 WO9935158-A1.
 15-JUL-1999.
 06-JAN-1999; 99WO-US000108.
 07-JAN-1998; 98US-0070657P.
 07-JAN-1998; 98US-0070658P.
 07-JAN-1998; 98US-0070682P.
 07-JAN-1998; 98US-0070704P.
 (HUMA-) HUMAN GENOME SCI INC.
 Ruben SM, Soppet DR, Ebner R, Lafleur DW, Ni J, Brewer LA;
 Olsen HS, Duan RD, Rosen CA;
 WPI; 1999-444190/37.
 P-FSDS; AAY38388, AAY38432, AAY38433, AAY38434.
 New isolated human genes and the secreted polypeptides they encode.
 Claim 1; Page 151; 227bp, English.
 This sequence represents a nucleic acid molecule which encodes a secreted
 human protein. The gene number is given in the descriptor line. The gene
 can be used to generate fusion proteins by linking to the gene to a human
 immunoglobulin Fc portion (e.g. AA206210) for increasing the stability of
 the fused protein as compared to the human protein only. The invention
 relates to 36 novel genes and their fragments (nucleic acid sequences:
 AA206219-206263; amino acid sequences AAY38386-Y38498) which are useful
 for preventing, treating or ameliorating medical conditions e.g. by

CC protein or gene therapy. Also, pathological conditions can be diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 36 polynucleotides, based on
CC which tissues they are most highly expressed in (see AA056219 for
CC described uses)

SQ Sequence 1224 BP; 243 A; 404 C; 346 G; 230 T; 0 U; 1 Other;

Query Match: 89.5%; Score 949.2; DB 2; Length 1224;
Best Local Similarity 98.9%; Pred. No. 2e-210;
Matches 1039; Conservative 0; Mismatches 4; Indels 8; Gaps 8;

```

OY 1 GACTCTGACCTTGAAGGGGCTTAAGGGGTGAGCAGGCTCTCCGAGACTCCCGCAGATG 60
DB 181 GACTCTGACCTTGAAGGGGCTTAAGGGGTGAGCAGGCTCTCCGAGACTCCCGCAGATG 240
OY 61 GAGAGAGCCATCTGTGCTCCCTGCTGAGGGGCTTCTGCTGCTGCTGCTGCTGCTGCTG 120
DB 241 GAGAGAGCCATCTGTGCTCCCTGCTGAGGGGCTTCTGCTGCTGCTGCTGCTGCTGCTG 300
OY 121 TTGATGGCAGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 301 TTGATGGCAGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
OY 181 GATAGTTGTATTC-AAGGGGATCCAGTTCAAAAGGCTTCAACAGGTTGGCCCGCTGGCC 239
DB 361 GATAGTTGTATTCAGAAAGGGGATCCAGTTCAAAAGGCTTCAACAGGTTGGCCCGCTGGCC 420
OY 240 ACCTGCTTACCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 299
DB 421 ACCTGCTTACCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
OY 300 CCCAAGATCCCGCAGCCCTTGGGGGCTTCCACCGGAGCCATCTTCCCGGGGATTC 359
DB 481 CCCAAGATCCCGCAGCCCTTGGGGGCTTCCACCGGAGCCATCTTCCCGGGGATTC 540
OY 360 TGATGGTGCACAGTGTGGGAGCTAGAGAAAGGAAAGGAGGCTTGAAGATGACAGA 419
DB 541 TGATGGTGCACAGTGTGGGAGCTAGAGAAAGGAAAGGAGGCTTGAAGATGACAGA 599
OY 420 TGAGATGAGAGAGCTATCACAAACCCAGGCTACCTGTGTGTGCTTCTGACAGCACCC 479
DB 600 TGA-GATGAGAGAGCTATCACAAACCCAGGCTACCTGTGTGTGCTTCTGACAGCACCC 658
OY 480 GGCCACTAGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
DB 659 GGCCACTAGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
OY 540 TGCCCTTCCAGTGAAGTCCATGATGATTAAGTGAACGTTCCGGAGAGGGGAGAGAGCC 599
DB 718 TGCCCTTCCAGTGAAGTCCATGATGATTAAGTGAACGTTCCGGAGAGGGGAGAGAGCC 777
OY 600 AGAAGCCTTCTGATGTGAGCGCCGAGATGTGATGTGTCCAGAGAACTGCATCTTGG 659
DB 778 AGAAGCCTTCTGATGTGAGCGCCGAGATGTGATGTGTCCAGAGAACTGCATCTTGG 837
OY 660 AGCGCTTAAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 719
DB 838 AGCGCTTAAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 897
OY 720 GGGGGCTCCAGATTGAGGAATCTGAGGAGTGAACCTGAGGGCCCTGAGAGCCGAGTGC 779
DB 898 GGGGGCTCCAGATTGAGGAATCTGAGGAGTGAACCTGAGGGCCCTGAGAGCCGAGTGC 957
OY 780 TGTCTCTGGAACAGGCTTGTGCTGAGACGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGG 839
DB 958 TGTCTCTGGAACAGGCTTGTGCTGAGACGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGG 1017
OY 840 TCTTCAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 899
DB 1018 TCTTCAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077

```

```

OY 900 ACTATATACATTGAGGGTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 959
DB 1078 ACTATATACATTGAGGGTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1135
OY 960 GCGTGAAGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1019
DB 1136 GCGTGAAGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1193
OY 1020 TGTGCTGTGTAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1050
DB 1194 TGTGCTGTGTAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1224

```

RESULT 3

ID ADQ22892 standard; DNA; 1488 BP.

AC ADQ22892;

DT 26-AUG-2004 (first entry)

DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5712.

XX soft tissue sarcoma; cyrostatic; gene therapy; vaccine; screening; human;

KW ds.

OS Homo sapiens.

PN WO2004048938-A2.

PD 10-JUN-2004.

PF 26-NOV-2003; 2003WO-US038193.

PR 26-NOV-2002; 2002US-0429739P.

PA (PROT-) PROTEIN DESIGN LABS INC.

PI Aziz N, Ginsburg WM, Zlotnick A;

XX WPI; 2004-441208/41.

PT Early detection of soft tissue sarcoma comprises determining expression

PT of a gene in a first soft tissue sample and a normal soft tissue sample

PT and comparing the gene expression, also useful in treating soft tissue

PT sarcoma.

XX Example 2; SEQ ID NO 5712; 210bp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma

XX which comprises obtaining a first soft tissue sample from an individual

XX and a normal soft tissue sample from the same or different individual,

XX determining the expression of a gene in both samples and comparing the

XX expression of the gene in both soft tissue samples, where a higher level

XX of protein expression in the first soft tissue sample indicates the

XX presence of soft tissue sarcoma. The method of the invention has

XX cyrostatic applications and may be useful for detecting soft tissue

XX sarcoma, possibly via gene therapy or vaccine production. The nucleic

XX acid sequences may be useful in diagnostic and screening applications.

XX The current sequence is that of a human soft tissue sarcoma-upregulated

XX DNA of the invention. The current sequence is not shown within the

XX specification per se but was submitted in CD format by the inventor.

SQ Sequence 1488 BP; 294 A; 443 C; 432 G; 319 T; 0 U; 0 Other;

Query Match: 87.9%; Score 932; DB 12; Length 1488;

Best Local Similarity 92.2%; Pred. No. 2.1e-206;

Matches 1029; Conservative 0; Mismatches 0; Indels 87; Gaps 1;

```

OY 1 GACTCTGACCTTGAAGGGGCTTAAGGGGTGAGCAGGCTCTCCGAGACTCCCGCAGATG 60
DB 22 GACTCTGACCTTGAAGGGGCTTAAGGGGTGAGCAGGCTCTCCGAGACTCCCGCAGATG 81

```

QY	6	AAGAGGGGCATCTGGTGTCCCTGGGAGCTGGGGGCTCTGGCTGGCCATCTGGGCGANTG	120
Db	82	GAGGAGGGGCATCTGGTGTCCCTGGGAGCTGGGGGCTCTGGCTGGCCATCTGGGCGANTG	141
QY	121	TTGATGGCACTGTGTGTGCACTGGCACACTGGCCAGAGCTGCTCTTACGACAGCAATCTTCA	180
Db	142	TTGATGGCACTGTGTGTGCACTGGCACACTGGCCAGAGCTGCTCTTACGACAGCAATCTTCA	201
QY	181	GATAGTTTGTATCCAAAGGGGCACTCCAGTTCAAAGGCTCTACAAGGTTGCCCCCTGGCCA	240
Db	202	GATAGTTTGTATCCAAAGGGGCACTCCAGTTCAAAGGCTCTACAAGGTTGCCCCCTGGCCA	261
QY	241	CCTGGCTTACCCCACTGTGCACTCTTACACCAACCCCTGAGCGACGCCAGAAGCTGTCCCATC	300
Db	262	CCTGGCTTACCCCACTGTGCACTCTTACACCAACCCCTGAGCGACGCCAGAAGCTGTCCCATC	321
QY	301	CCAAAGATCCCGCAGACCCCTTGGGGGCTCTCCACCCGACGCAATCTTCCCGGCGGGATTTCT	360
Db	322	CCAAAGATCCCGCAGACCCCTTGGGGGCTCTCCACCCGACGCAATCTTCCCGGCGGGATTTCT	381
QY	361	GATGGTGGCAACAGGTGTGGCGAGCTAACGAGACGAG-----	396
Db	382	GATGGTGGCAACAGGTGTGGCGAGCTAACGAGACGAGGGTGGCTTGGGCAATCCGAGGTGCC	441
QY	397	-----	396
Db	442	CAGGCTGGGTGGGAGAGTCTGGGGGTCCGTCCTGGACTAGAGGCTGACCCCTGTGTGTTACCC	501
QY	397	---GAAACCAAGCTGTGAGAGATGCAATATGAGATGAGAGACGATATTCACAAACCAAGGCTAC	453
Db	502	CCAGAAACCAAGCTGTGAGAGATGCAATATGAGATGAGAGACGATATTCACAAACCAAGGCTAC	561
QY	454	CTGGTGTGTCTTCTCTGACAGCACCCCGGCACTTACACCTGTGCCCCATCAGCTCCTGCA	513
Db	562	CTGGTGTGTCTTCTCTGACAGCACCCCGGCACTTACACCTGTGCCCCATCAGCTCCTGCA	621
QY	514	CTCAGCACCCCTGGCATCCGAGACAGTGTCTTCTCCATGAGAGTCCATGTGATTTACGTG	573
Db	622	CTCAGCACCCCTGGCATCCGAGACAGTGTCTTCTCCATGAGAGTCCATGTGATTTACGTG	681
QY	574	AACGTTCCGGAGAGGGGGAGAGGCCCAAGAAAGGCTCTGCGATGAGGACCCGGGAGATTATGG	633
Db	682	AACGTTCCGGAGAGGGGGAGAGGCCCAAGAAAGGCTCTGCGATGAGGACCCGGGAGATTATGG	741
QY	634	AATGTGTCCCAAGAACTGTCATCTGTAGCGGCTTAAGACTGAGCCTTGCCTCTGAGTTCC	693
Db	742	AATGTGTCCCAAGAACTGTCATCTGTAGCGGCTTAAGACTGAGCCTTGCCTCTGAGTTCC	801
QY	694	CAGGAGGCAAGAGAACTGAGAGAAAGGGGGCTCCAGATTACGAGAAATTGCAAGAGACTG	753
Db	802	CAGGAGGCAAGAGAACTGAGAGAAAGGGGGCTCCAGATTACGAGAAATTGCAAGAGACTG	861
QY	754	AACCTAGAGGGCTGTGTGAGAGCCGAGTCTGTCTTGAAACAAGGCTGTGGAGCCGCTGAG	813
Db	862	AACCTAGAGGGCTGTGTGAGAGCCGAGTCTGTCTTGAAACAAGGCTGTGGAGCCGCTGAG	921
QY	814	CTGGGCACTGGAAGTGTGCTTGTGGGGTCTCTCACTATGCGTCTGTGCCCTTGTCTCAAGCTG	873
Db	922	CTGGGCACTGGAAGTGTGCTTGTGGGGTCTCTCACTATGCGTCTGTGCCCTTGTCTCAAGCTG	981
QY	874	ACAAAGGCTGAGAAATCCCGCCGTAATTATTTACCTTGGGGGTGGGCTGTGTGCC	933
Db	982	ACAAAGGCTGAGAAATCCCGCCGTAATTATTTACCTTGGGGGTGGGCTGTGTGCC	1044
QY	934	CCGAAGGCTGTGCACTTCTGACGAGCCTGTGAATGACCTGCCCTGGGCCCAAGCCTTAC	993
Db	1042	CCGAAGGCTGTGCACTTCTGACGAGCCTGTGAATGACCTGCCCTGGGCCCAAGCCTTAC	1100
QY	994	TCTGTGTATATGAATTAAGAGCCTGCGTGTGTCTGTG 1029	
Db	1102	TCTGTGTATATGAATTAAGAGCCTGCGTGTGTCTGTG 1137	

ID	AA899073	standard; DNA, 1461 BP.
AC	AA89073;	
XX		
DT	14-SEP-1999	(first entry)
DE		Human LAT alternative splice variant nucleotide sequence.
KW	LAT, tyrosine kinase; linker for activation of T cell; TCR; human; T-cell receptor; TCR signalling pathway; neoplasia; inflammation; hypersensitivity; allergy; microbial infection; genetic disease; autoimmune disease; graft rejection; modulator; variant; ss.	
OS	Homo sapiens.	
PN	WO932627-A2.	
XX		
PD	01-JUL-1999.	
PF	23-DEC-1998;	98WO-US027400.
PR	23-DEC-1997;	97US-0068690P.
XX		
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
PI	Samelson LE, Zhang W;	
DR	WPI, 1999-418926/35.	
XX		
PT	Linker for activation of T cell protein used to, e.g. screen for modulators of T cell signaling.	
PS		
XX	Disclosure; Fig 7B; 125pp; English.	
CC	The invention relates to a protein tyrosine kinase substrate LAT (linker for activation of T cells) protein. Modulation of interaction between LAT and the T-cell receptor (TCR) affects the TCR signalling pathway. LAT is a substrate for tyrosine kinases and becomes phosphorylated after TCR engagement, resulting in recruitment of other signalling molecules. LAT is used to identify and test (antagonists of tyrosine kinase signalling pathways, i.e. modulation of interaction between tyrosine kinase substrates and intracellular ligands or between these ligands and other members of the pathway, including identification of downstream signalling proteins, particularly in immune system cells. These modulators are potentially useful as drugs and diagnostic agents, particularly for diseases that involve undesirable cell proliferation, differentiation, growth or T cell energy, e.g. neoplasia, inflammation, hypersensitivity/allergy, microbial infection, metabolic, genetic or autoimmune diseases, graft rejection. LAT is also used to generate specific antibodies, used for detection of LAT. Nucleic acid that encodes LAT, or its fragments, are used to identify homologous sequences in other species; to detect the LAT gene and as sources of antisense therapeutics. Modulators of LAT are potentially more specific and less toxic than known immunosuppressants such as cyclosporin. The present sequence represents the nucleotide sequence of an alternative splice variant of human LAT	
SQ	Sequence 1461 BP; 269 A; 443 C; 432 G; 317 T; 0 U; 0 Other;	
Query Match	86.9%; Score 921; DB 2; Length 1461;	
Best Local Similarity	92.1%; Pred. No. 7.4e-204;	
Matches 1029,	Conservative 0; Mismatches 0; Indels 88; Gaps 2;	
OY	1 GACTTGGCCCTTGAGGCGCTTAGGGGTGCACGCCCTTCGCGAGCTCCTTGCAATG	60
DB	22 GACTTGGCCCTTGAGGCGCTTAGGGGTGCACGCCCTTCGCGAGCTCCTTGCAATG	81
OY	61 GAGAGGAGCATCTCGTGCCTCCGTCGTGGGGCTCTGCTGCTGCCATCTGGACCATT	120
DB	82 GAGGAGGCGATCTGTGTCTCCCTCGCTGCGGGCTCTGCTGCTGCCATCTGGACCATT	141
OY	121 TTGAATGAGCACTGTGTGTGTCACATGCAAGAATCGAGGCTCCTTAACAAGACATCTCTCA	180


```
Db 142 TTGATGCACTGTGTGTGCACTGCCACAGACTGCCAGGCTCTTACAGACAGACATCTCTCA 201
QY 181 GATTAGTTTATTCACAGGGGCAATCCAGTTTCAAAAGCCCTTCAACCGTTGCCCTTGCCA 240
Db 202 GATAGTTTATTCACAGGGGCAATCCAGTTTCAAAAGCCCTTCAACCGTTGCCCTTGCCA 261
QY 241 CTTGCTTACCACTGTGTGTGCACTGCCACAGGCTCTTACAGACAGACATCTCTCTCA 300
Db 262 CTTGCTTACCACTGTGTGTGCACTGCCACAGGCTCTTACAGACAGACATCTCTCTCA 321
QY 301 CCAAGATCCCGGAGCCCTTGGGGGCTCCACCGGAGCCCATC-TTCCCGGGGGATTC 359
Db 322 CCAAGATCCCGGAGCCCTTGGGGGCTCCACCGGAGCCCATCTTCCCGGGGGATTC 381
QY 360 TGTAGTGTGCAACAGTGTGGCCGACGCTACGAAACGAG----- 396
Db 382 TGTAGTGTGCAACAGTGTGGCCGACGCTACGAAACGAGGGTGCCTTGGGATCCGAGGTGC 441
QY 397 ----- 396
Db 442 CCAAGCTGGGTGGGAGTCTGGGGTCCGTCTGGAATGAGCTGACCCCTGTGTCTTACC 501
QY 397 -----GAAACAGCTGTGAGATGACAGATGAGATGAGAGACGACTATACAGACCCAGGCTA 452
Db 502 CCAAGATCCCGGAGCCCTTGGGGGCTCCACCGGAGCCCATCTTCCCGGGGGATTC 561
QY 453 CTTGCTTACCACTGTGTGTGCACTGCCACAGGCTCTTACAGACAGACATCTCTCTCA 512
Db 562 CTTGCTTACCACTGTGTGTGCACTGCCACAGGCTCTTACAGACAGACATCTCTCTCA 621
QY 513 ACTGACACACCCCTGGGATCCGAGACAGTGTCTTCCATGAGGATGATGATGATGATGATGAT 572
Db 622 ACTGACACACCCCTGGGATCCGAGACAGTGTCTTCCATGAGGATGATGATGATGATGATGAT 681
QY 573 GAACTGTCGGAAGAGGGGAGAGCGCAGAAAGGCTCTGATGAGAGCGGGAGATGAT 632
Db 682 GAACTGTCGGAAGAGGGGAGAGCGCAGAAAGGCTCTGATGAGAGCGGGAGATGAT 741
QY 633 GAACTGTCGGAAGAGGGGAGAGCGCAGAAAGGCTCTGATGAGAGCGGGAGATGAT 692
Db 742 GAACTGTCGGAAGAGGGGAGAGCGCAGAAAGGCTCTGATGAGAGCGGGAGATGAT 801
QY 693 CAGAGAGGAGAGAGTGTGAGAGAGAGGGGCTCCAGATTGAGAAATCTGAGAGAGCT 752
Db 802 CAGAGAGGAGAGAGTGTGAGAGAGAGGGGCTCCAGATTGAGAAATCTGAGAGAGCT 861
QY 753 GAACTGAGGGCTGTGAGAGGGGAGTGTCTGAGAAACGAGGCTTCTGAGAGCGGCTGA 812
Db 862 GAACTGAGGGCTGTGAGAGGGGAGTGTCTGAGAAACGAGGCTTCTGAGAGCGGCTGA 921
QY 813 GCTGAGGAGCTGTGAGAGGGGCTGTGAGAGGGGCTTCTGAGAGGGGCTTCTGAGAGCT 872
Db 922 GCTGAGGAGCTGTGAGAGGGGCTGTGAGAGGGGCTTCTGAGAGGGGCTTCTGAGAGCT 981
QY 873 GAACTGAGGGCTGTGAGAGGGGCTGTGAGAGGGGCTTCTGAGAGGGGCTTCTGAGAGCT 932
Db 982 GAACTGAGGGCTGTGAGAGGGGCTGTGAGAGGGGCTTCTGAGAGGGGCTTCTGAGAGCT 1041
QY 933 CCGAAGAGCTGTGAGAGGGGCTGTGAGAGGGGCTTCTGAGAGGGGCTTCTGAGAGCT 992
Db 1042 CCGAAGAGCTGTGAGAGGGGCTGTGAGAGGGGCTTCTGAGAGGGGCTTCTGAGAGCT 1101
QY 993 CTCTGTGTATATGAATTAAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1029
Db 1102 CTCTGTGTATATGAATTAAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1138
```

RESULT 5
AD018376/c
ID AD018376 standard; DNA; 518 BP.
XX
AC AD018376;

```
XX 26-AUG-2004 (first entry)
DT Human soft tissue sarcoma-upregulated DNA - SEQ ID 1195.
XX soft tissue sarcoma; cytosolic; gene therapy; vaccine; screening; human;
XX ds.
XX Homo sapiens.
XX WO2004048938-A2.
XX 10-JUN-2004.
XX 26-NOV-2003; 2003WO-US038193.
XX 26-NOV-2002; 2002US-0429739P.
XX (PROT-) PROTEIN DESIGN LABS INC.
XX Aziz N, Ginsburg WM, Zlotnick A;
XX WPI; 2004-441208/41.
XX Early detection of soft tissue sarcoma comprises determining expression
XX of a gene in a first soft tissue sample and a normal soft tissue sample
XX and comparing the gene expression, also useful in treating soft tissue
XX sarcoma.
XX Example 2; SEQ ID NO 1195; 210pp; English.
XX
XX The invention relates to a novel method for detecting soft tissue sarcoma
XX which comprises obtaining a first soft tissue sample from an individual
XX and a normal soft tissue sample from the same or different individual,
XX determining the expression of a gene in both samples and comparing the
XX expression of the gene in both soft tissue samples, where a higher level
XX of protein expression in the first soft tissue sample indicates the
XX presence of soft tissue sarcoma. The method of the invention has
XX cytosolic applications and may be useful for detecting soft tissue
XX sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX acid sequences may be useful in diagnostic and screening applications.
XX The current sequence is that of a human soft tissue sarcoma-upregulated
XX DNA of the invention. The current sequence is not shown within the
XX specification per se but was submitted in CD format by the inventor.
XX
XX Sequence 518 BP; 116 A; 146 C; 149 G; 107 T; 0 U; 0 Other;
XX
XX Query Match 42.8%; Score 454; DB 12; Length 518;
XX Best Local Similarity 92.3%; Pred No. 1.5e-95;
XX Matches 478; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 520 ACCCTGGGATCCGAGACAGTGTCTTCCATGAGAGTCCATGATGATGATGATGATGATGAT 579
Db 518 ACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 459
QY 580 CCGAGAGGAGGAGAGAGCGCAGAAAGGCTCTGATGAGAGCGGGAGATGATGATGATGAT 639
Db 458 CCGAGAGGAGGAGAGAGCGCAGAAAGGCTCTGATGAGAGCGGGAGATGATGATGATGAT 399
QY 640 TCCCAAGAACTGATCTGAGAGGGGCTTAAGATGAGAGCGGCTGAGAGCGGCTGAGAGCG 699
Db 398 TCCCAAGAACTGATCTGAGAGGGGCTTAAGATGAGAGCGGCTGAGAGCGGCTGAGAGCG 339
QY 700 GCAGAGAGAGTGTGAGAGAGGGGCTCCAGATTACGAAATCTGAGAGCTGAGAGCTGA 759
Db 338 GCAGAGAGAGTGTGAGAGAGGGGCTCCAGATTACGAAATCTGAGAGCTGAGAGCTGA 279
QY 760 GGGGCTGTGAGAGCGGAGTGTCTTCTGAGAAACGAGGCTTCTGAGAGCGGCTGAGAGCG 819
Db 278 GGGGCTGTGAGAGCGGAGTGTCTTCTGAGAAACGAGGCTTCTGAGAGCGGCTGAGAGCG 219
QY 820 AGCTGAAGTGTGCTGTGGGGTCTTCAATGAGGCTCTGAGAGCGGCTTCTGAGAGCGGCTGA 879
```


PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX SameLson LE, Zhang W;

XX WPI, 1999-418926/35.

DR P-PSDB, AAY27121.

PT Linker for activation of T cell protein used to, e.g. screen for
PT modulators of T cell signaling.

PS Example 2, Fig 7C, 125pp; English.

XX The invention relates to a protein tyrosine kinase substrate LAT (linker
XX for activation of T cells) protein. Modulation of interaction between LAT
XX and the T-cell receptor (TCR) affects the TCR signaling pathway. LAT is
XX a substrate for tyrosine kinases and becomes phosphorylated after TCR
XX engagement, resulting in recruitment of other signaling molecules. LAT
XX is used to identify and test (ant)agonists of tyrosine kinase signaling
XX pathways, i.e. modulation of interaction between tyrosine kinase
XX substrates and intracellular ligands or between these ligands and other
XX members of the pathway, including identification of downstream signalling
XX proteins, particularly in immune system cells. These modulators are
XX potentially useful as drugs and diagnostic agents, particularly for
XX diseases that involve undesirable cell proliferation, differentiation,
XX growth or T cell anergy, e.g. neoplasia, inflammation, hypersensitivity/
XX allergy, microbial infection, metabolic, genetic or autoimmune diseases,
XX graft rejection. LAT is also used to generate specific antibodies, used
XX for detection of LAT. Nucleic acid that encodes LAT, or its fragments,
XX are used to identify homologous sequences in other species; to detect the
XX LAT gene and as sources of antisense therapeutics. Modulators of LAT are
XX potentially more specific and less toxic than known immunosuppressants
XX such as cyclosporin. The present sequence represents the nucleotide
XX sequence of murine LAT

SQ Sequence 1260 BP; 282 A; 359 C; 360 G; 259 T; 0 U; 0 Other;

Query Match 37.8%; Score 400.4; DB 2; Length 1260;

Best Local Similarity 67.6%; Pred. No. 5,5e-83;

Matches 698; Conservative 0; Mismatches 291; Indels 43; Gaps 8;

QY 26 GTGACGACGAGCTGCTCCGAGCTCCCTGAGATGAGAGGACCATCTGTCCTCGG 85
DB 260 GTGACGACGAGGCTGTGAGAGCTGTCTGACATGGAAGACAGCCTTGAAGCCGGTGG 319
QY 86 TGCTGGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 142
DB 320 GCTGGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 379
QY 143 GCCACAGATGCGACGAGCTCTGACAGACATCTCAGATGTTGTTATCCAGAGGGGCA 202
DB 380 GCCGAGATGCTGCACTCTCTATGACAGACATCTTCAACAGAGTTGTAACCAAGAGCA 439
QY 203 TCCAGTTCAAGGCGCTGACAGGCTTGCCTGCGACCTG-----CTTACCAACCTG 256
DB 440 TCCATCATGACGACCTCAATTAACCGTCCCGAAGACCTGCTGTTCTTCACTCCCTAG 499
QY 257 TCACTCTCTAAGCCAGCTGAGCCAGCCAGACCTGCTCCCATCCAGATCCCGGACG 316
DB 500 TCATCTCTCTTCCACCCCTGAGGACGACGACCTGCTCCCATCCCGAATCCCAACAGC 559
QY 317 CCTTGGGGGGCTCCACGCGAGCGCATCTTCCGGGGGATCTGTAGTGTGCAACAGTG 376
DB 560 CCTTGGGGGGTCCATGCGATGCACTCTTCCAGAGATTCAGTGTGCAACAGTG 619
QY 377 TGGCAGGCTACGAGAACGAGAACGAGCTGTGAGATGACATGAGATGAGAGCACT 436
DB 620 TGGCAGGCTACGAGAACGAGAACGAGCTGTGAGATGAGATGAGATGAGAGCACT 679
QY 437 ATCAACACCC--AGGCTACCTGTGTGTGTCTTCTGACAGCACCCCGGACATAGACTG 493
DB 680 ACGACTATCTCCAAACGCTACCTAGTGTGTGTCTTCTGACAGTGTGTGTGTGTGTGTGTGT 739
QY 494 CTGCCCATCAGCTCTGCACTCAGAACCCCTGGAGATCCGAGACAGTGCTTCTCCATGG 553

DB 740 TTGTCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 799
QY 554 AGTCATTGATGATTTACGTAAGACCTTCCGAGAGCGGGAGAGCCGACAGGCTCTCTGG 613
DB 800 AGTCATTGATGATTTACGTAAGACCTTCCGAGAGCGGGAGAGCCGACAGGCTCTCTGG 859
QY 614 ATGGAGCGCGGAGATGTAATGATGCTCCAGGAAGCTGATCTGAGCGGCTTAAGACTG 673
DB 860 ATGGAGCGCGGAGATGTAATGATGCTCCAGGAAGCTGATCTGAGCGGCTTAAGACTG 916
QY 674 AGCTGCGCGCTCTGAGTTTCCAG 733
DB 917 AGCTGCGCGCTCTGAGTTTCCAG 976
QY 734 ACGAATATGAG 793
DB 977 GAG 1036
QY 794 GCTTCCGCGAG 853
DB 1037 TCTGTCTG-----AACTGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1088
QY 854 CTTGCTG 913
DB 1089 CTTGCTG 1136
QY 914 TGGGCTTGGGCTG 973
DB 1137 TGGGCTTGGGCTG 1194
QY 974 TGGGCTTGGGCTG 1033
DB 1195 CTGTTTCCAGCCAGAC-----ATGTCATGGAATTAAGCCATGTGACATTAAGAGAGAGAGAGAGAGAG 1248
QY 1034 AAAAAAAAAAAAAA 1045
DB 1249 AAAAAAAAAAAAAA 1260

RESULT 8
ID AAK57405
AKS7405 strand; cDNA; 312 BP.
XX
AC AAK57405;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human Immune/haematopoietic antigen encoding cDNA SEQ ID NO:2465.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ss.
OS Homo sapiens.
PN W0200157182-A2.
PD 09-AUG-2001.
PF 17-JAN-2001; 2001WO-US001354.
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.

CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 312 BP; 64 A; 113 C; 70 G; 64 T; 0 U; 1 Other;

Query Match 7.8%; Score 83; DB 4; Length 312;
Best Local Similarity 89.3%; Pred. No. 1.8e-09;
Matches 134; Conservative 0; Mismatches 11; Indels 5; Gaps 4;

QY 303 AAGATCCCGCGGCGCCCTTTGGGGGCTCCACCGGAGCCATCTTCCCGCGGGA-TTCTG 361
Db 163 AAGATCCCGCGGCGCCCTTTGGGGGCTCCACCGGAGCCATCTTCCCGCGGATTCTG 222

QY 362 ATGGTGCACACAGTGTGGGAGCT-ACGAGACGAGAACGACCTGTG-AGATGCAGA 419
Db 223 ATGGTGCACACAGTGTGGGAGCTAACGAGAACGAGAACGACCTGTGAGAGATGCAGA 282

QY 420 TGA--GGATGAGAGAGACTATCAACCA 447
Db 283 TGAAGATGAGAGAGAGACTATCAACCA 312

RESULT 9
AAK57403
ID AAK57403 standard; cDNA; 519 BP.
XX
AC AAK57403;
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:2463.
XX
KM Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KM cytoskeletal; gene therapy; vaccine; metastasis; ss.
OS Homo sapiens.
XX
XX MO200157182-A2.
PN 09-AUG-2001.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216890P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227099P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0235837P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-0239393P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.

PR	08-NOV-2000	2000US-0246527P
PR	08-NOV-2000	2000US-0246528P
PR	08-NOV-2000	2000US-0246532P
PR	08-NOV-2000	2000US-0246609P
PR	08-NOV-2000	2000US-0246610P
PR	08-NOV-2000	2000US-0246611P
PR	08-NOV-2000	2000US-0246613P
PR	17-NOV-2000	2000US-0249207P
PR	17-NOV-2000	2000US-0249208P
PR	17-NOV-2000	2000US-0249209P
PR	17-NOV-2000	2000US-0249210P
PR	17-NOV-2000	2000US-0249211P
PR	17-NOV-2000	2000US-0249212P
PR	17-NOV-2000	2000US-0249213P
PR	17-NOV-2000	2000US-0249214P
PR	17-NOV-2000	2000US-0249215P
PR	17-NOV-2000	2000US-0249216P
PR	17-NOV-2000	2000US-0249217P
PR	17-NOV-2000	2000US-0249218P
PR	17-NOV-2000	2000US-0249244P
PR	17-NOV-2000	2000US-0249245P
PR	17-NOV-2000	2000US-0249264P
PR	17-NOV-2000	2000US-0249265P
PR	17-NOV-2000	2000US-0249297P
PR	17-NOV-2000	2000US-0249299P
PR	17-NOV-2000	2000US-0249300P
PR	01-DEC-2000	2000US-0250160P
PR	01-DEC-2000	2000US-0250391P
PR	05-DEC-2000	2000US-0251030P
PR	05-DEC-2000	2000US-0251988P
PR	05-DEC-2000	2000US-0256719P
PR	06-DEC-2000	2000US-0251479P
PR	08-DEC-2000	2000US-0251566P
PR	08-DEC-2000	2000US-0251688P
PR	08-DEC-2000	2000US-0251699P
PR	08-DEC-2000	2000US-0251899P
PR	11-DEC-2000	2000US-0251990P
PR	05-JAN-2001	2000US-0254097P
PR	05-JAN-2001	2001US-0259678P

PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
DR
XX P-PsDB; AAM64622.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis
PT
XX
PS Claim 1; SEQ ID NO 2463; 3071pp + Sequence Listing; English.

CC AAK65495 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM89170 to AAM91921. (I) have cytotoxic
CC activity, and can be used in gene therapy and vaccine production. (II)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK64950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX

SQ Sequence 519 BP; 80 A; 156 C; 161 G; 118 T; 0 U; 4 Other;

Query Match 6.4%, Score 68; DB 4; Length 519;

	Best Local Similarity	80.0%;	Pred.	No. 6.4e-06;	Mismatches	20;	Indels	0;	Gaps	0;
	Matches	80;	Conservative	0;	Mismatches	20;	Indels	0;	Gaps	0;
Oy	544	TTCTTCATGAGAGTCCATTGATGATTACGTGAACGTTCCCGAGAGCGGGAGAGGCCAGAA	603							
Db	247	TCCACAGTGAGTCCATTGATGATTACGTGAACGTTCCCGAGAGCGGGAGAGGCCAGAA	306							
Oy	604	GCGTCTGTGATGAGCGCGGAGATATGTATGTATGTCCTCC	643							
Db	307	GCGTCTGTGATGAGTAACCGGTCCTTTCGCGTCCCTCCTCC	346							

RESULT	10
ID	ADA19460
	ADA19460 standard; DNA; 6307 BP.
XX	
AC	ADA19460;
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	Mouse LAT Y136F gene SEQ ID NO:1.

KW mutant; linker for activation of T-cells; LAT, TH2 cell differentiation
 KW allergy/ asthma; eosinophilia; TH2 cell deregulation; antiasthmatic;
 KW antiallergic; gene; ds.
 XX
 Mus musculus.
 OS

PN MO2003068968-A2.
XX
PD 21-AUG-2003.
XX
PF 14-FEB-2003; 2003MO-IB001044.

PR 14-FEB-2002; 2002US-0356136P.
PR 11-MAR-2002; 2002EP-00290610.
XX
XX (INRM) INSRM INST NAT SANTE & RECH MEDICALE.
XX

XX New non-human animal (e.g. mouse) comprising a mutated linker for
 DR Activation of T-cells (Lkr) gene coding for a mutant Lkr protein, useful
 XX as a model of allergy, asthma, eosinophilia or a disease associated with
 PT Tkr cell deregulation.
 WI MPI; 2003-679639/64.

PS Claim 30; Page 46-50; 52pp; English

The present invention describes a non-human animal having a mutated LAT linker for activation of T-cells (LAT) gene coding for a mutant LAT protein, where the mutant LAT protein leads to an exaggerated TH2 cell differentiation. Also described: (1) a germ cell or somatic cell from a non-human animal or its progeny containing the mutated LAT gene; (2) a screening for a drug for treating allergy, asthma and/or disease associated with TH2 cell deregulation, or for drugs that regulate the activity of TH2 cells; (3) producing a pharmaceutical composition for treating a disease associated with deregulated TH2 cells activity; (4) a bioreactor for a large scale production of human IGE antibodies comprising an animal; (5) producing human IGE antibodies; (6) a B-cell hybridoma obtained by the above method; (7) a mutated mouse gene coding for a mutant LAT protein, the sequence of which corresponds to a wild type sequence and contains a single mutation of the tyrosine Y136 or a composite mutation of the three distal tyrosine residues; (8) a diagnostic method for asthma, allergy, eosinophilia and/or TH2 cells deregulation; (9) a diagnostic kit for asthma, allergy, eosinophilia and/or TH2 cells deregulation, comprising oligonucleotide probes for the detection of a mutated LAT gene coding for a mutant LAT protein containing a single mutation of the tyrosine Y136 or a composite mutation of the three distal tyrosine residues; and (9) a non-human animal resulting from the breeding of a non-human animal expressing humanized IGE with the non-human animal described above. Mutant LAT has

CC antiasthmatic and antiallergic activities. The animal is useful as a
CC model of allergy, asthma, eosinophilia and/or a disease associated with
CC TH2 cell deregulation. The methods may be used for screening drugs that
CC may be utilized for treating asthma or allergy or for diagnosing the
CC above diseases. The present sequence represents a mutated mouse LAT gene
CC which encodes a LAT Y136F protein.

XX Sequence 6307 BP; 1219 A; 1676 C; 1556 G; 1856 T; 0 U; 0 Other;

Query Match 5.2%; Score 54.6; DB 9; Length 6307;

Best Local Similarity 71.3%; Pred. No. 0.017;

Matches 72; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Oy 457 GTGGTCTCTCTGACAGACACCCGACCTAGACACTGCTCCCATCAGCTCTGCACTC 516

Db 2048 GTGGTCTCTCTGACAGAGTGTCTCTGCTCCGCTCTGCTCTCTCTCTCTGCTCT 2107

Oy 517 AGCACCCCTGGCATCCGACAGATGCTTCTTCATGAGACT 557

Db 2108 AGCAACCTGACCTTGAGAGAGTGGCTTCTCTGGAGACTC 2148

RESULT 11

ADA71938 standard; DNA; 2000 BP.

ADA71938;

20-NOV-2003 (first entry)

Rice gene, SEQ ID 5263.

Plant; bacterial infection; fungal infection; viral infection; rice;

gene; db.

Oryza sativa.

MO2003000898-A1.

03-JAN-2003.

22-JUN-2001; 2001WO-1B001105.

(SYGN) SYNGENTA PARTICIPATIONS AG.

Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

WPI, 2003-175290/17.

Identifying at least one gene involved in plant resistance or response to

pathogenic infection for conferring resistance or tolerance to a plant to

bacterial, fungal or viral infection by determining or detecting plant

gene expression.

Claim 27; SEQ ID NO 5263; 8999P; English.

The present invention relates to a method (M1) for identifying genes

involved in plant resistance or response to pathogenic infection. M1

comprises identifying a gene whose expression is significantly altered in

the incompatible interaction of plant gene expression relative to

expression of the gene in an uninfected plant, in a mutant plant that

does not express a gene associated with response to pathogenic infection,

or in a corresponding incompatible or compatible interaction. (M1) is

useful for conferring resistance to resistance or tolerance to a plant to

Best Local Similarity 10.1%; Pred. No. 0.026;

Matches 76; Conservative 355; Mismatches 308; Indels 11; Gaps 4;

Oy 313 CAGCCCTTGGGGGCTCCACCGACGCACTTCCCGGGGATTTTCATGTGCTCCAC 372

Db 23 MAGMMSCARMSSSRKMSKRSKRSKRSKRSKRSKRSKRSKRSKRSKRSKRSKRS 82

Oy 373 AGTGTGGAGAGCTAG 432

Db 83 GSYSGKMKKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRY 142

Oy 433 GACTATCA-----CAACCCAGCTACTGCTGTGCTCTCTTCTGACAGACACCCGACACTA 487

Db 143 SYMMVCYARAGGSCCRKSKSKSGSWKTCRARGSGSGSSGAKYSGSGSKMMSSCGR 202

Oy 488 GCACCTGCTGCCCCATCAGCTCTGCACTAG--CAACCTGCACTCCAGACAGAGCTCT 545

Db 203 SCGGRSASVSRYYGTSRKYGTGKMTYSSASCMAYMTYSMASSSYWCRSKRSRM 262

Oy 546 CTCGATGAGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 605

Db 263 WDMRKMRSRSGYMSYMSYMSYMSYMSYMSYMSYMSYMSYMSYMSYMSYMSYMS 322

Oy 606 GTCTCTGATGAGACCGGAGATGATGATGATGATGATGATGATGATGATGATGATG 665

Db 323 AMYKMTYMYRKYKMYKMYKMYKMYKMYKMYKMYKMYKMYKMYKMYKMYKMYKMY 382

Oy 666 TAAAGCTAGAGCTGCGCCCTGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 725

Db 383 SRKRKRYKMYRKYKMYRKYKMYRKYKMYRKYKMYRKYKMYRKYKMYRKYKMYR 442

Oy 726 TCCAGATTACGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 785

Db 443 RWSKRMASKYMMSRYMRYMRYMRYMRYMRYMRYMRYMRYMRYMRYMRYMRYM 502

Oy 786 GGAACGAGCTGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 845

Db 503 RRMGRMYRMRYMRYMRYMRYMRYMRYMRYMRYMRYMRYMRYMRYMRYMRYMRY 562

Oy 846 CATGGCGTC--CTGCCCTTGTCTGAGCTGACCAAGCTGAGAAATCCCTGTAAT 902

Db 563 YMYKMMYMYRKYKMYRKYKMYRKYKMYRKYKMYRKYKMYRKYKMYRKYKMYR 622

Oy 903 TATTTACATTGGGCTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 962

Db 623 CASKRSASRYAMMGMGMSGMSRMSKSYTCYWRMGMSKSYTCYWRMGMSKSYTC 682

Oy 963 TG-AGAAAGAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021

Db 683 WRIRYRAMCMYMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 742

Oy 1022 TGTCTGTGAAAAAATAAAAAAAAAAAAAAAAAAAAAA 1051

Db 743 YRSMYMYSMWMAKTWKMRYATRMWMMW 772

RESULT 12

AAC78190 standard; cDNA; 2968 BP.

AAC78190;

08-FEB-2001 (first entry)

Human cancer associated gene sequence SEQ ID NO:584.

Human; cancer associated gene; cancer antigen; detection; cancer;

diagnosis; cytostatic; proliferative; vulnery; immunomodulator;

antidiabetic; antiasthmatic; antineoplastic; antiallergic; antiviral;

antitumor; antitumor; antitumor; antitumor; antitumor; antitumor;

dermatological; neoplastic; thrombolytic; coagulant; neoplastic;

vasotropic; antiproliferative; angiogenic; gene therapy; inflammation;

immune disorder; haematopoietic cell disorder; autoimmune disorder;

KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening; ss.
XX Homo sapiens.
OS
PN WO200055350-A1.
XX
XX 21-SEP-2000.
PD
XX 08-MAR-2000; 2000WO-US005882.
PF
XX 12-MAR-1999; 99US-0124270P.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM;
PI
XX WPI, 2000-587533/55.
DR P-PSDB; AAB43981.
XX
XX Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer.
PS
XX Claim 1; Page 1106-1107; 2352pp; English.

CC AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnery; immunomodulator;
CC antidiabetic; antiaesthetic; antihemetic; antibacterial; antiviral;
CC antiinflammatory; neuroprotective; cardiac; thrombolytic; coagulant;
CC dermatologic; vasoregic; antipsoriatic and angiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention
XX
XX

SQ Sequence 2968 BP; 678 A; 817 C; 797 G; 671 T; 0 U; 5 Other;

Query Match 4.7%; Score 49.6; DB 3; Length 2968;
Best Local Similarity 53.6%; Pred. No. 0.2;
Matches 103; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 869 GCGTGAACAAGCCTGAGAAATCCCGTAATTAATTAATCACTTTGGGCTTGGCTGT 928
DB 2733 GCCCGAAGAGCCTCTTCGATAGGAGATGATTTGCAATTAATTAAGCTCACCTTTT 2792
QY 929 GTCCCCGGAAGCCTGACCTTCTGACGAGCCTGAGAAATGACCTGCCCTGGCCAGC 988
DB 2793 TTCTTCCCTCTCTGCTGCTGCTGCTGCAATTAATGATTTGTTCTTCCCTATGGGATC 2852
QY CCTACTCTGTGTATAGATTAAGAGCGCTGCTGTGTGTGGAATAAAAAAAAAAAAAA 1048
DB 2853 CATCTGTTTGTATAACAATAAAGGCTCTGAGGAGATGTAAAAA 2912
QY 1049 AAAAAAAAAAAAAA 1060
DB 2913 AAAAAAAAAAAAAA 2924

RESULT 13
AAH34408

ID AAH34408 standard; cDNA; 2968 BP.
XX
XX AAH34408;
XX
XX 03-SEP-2001 (first entry)
DT
XX
XX Human colon cancer antigen encoding cDNA SEQ ID NO:1490.
DE
XX
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.
XX
XX Homo sapiens.
OS
PN WO200122920-A2.
XX
XX 05-APR-2001.
PD
XX 28-SEP-2000; 2000WO-US026524.
PF
XX 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
PI
XX WPI; 2001-235357/24.
DR P-PSDB; AAG75003.
XX

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX
XX Claim 1; Page 3154-3155; 9803pp; English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patients own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX
XX

SQ Sequence 2968 BP; 678 A; 817 C; 798 G; 671 T; 0 U; 4 Other;

Query Match 4.7%; Score 49.6; DB 4; Length 2968;
Best Local Similarity 53.6%; Pred. No. 0.2;
Matches 103; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 869 GCGTGAACAAGCCTGAGAAATCCCGTAATTAATTAATCACTTTGGGCTTGGCTGT 928
DB 2733 GCCCGAAGAGCCTCTTCGATAGGAGATGATTTGCAATTAATTAAGCTCACCTTTT 2792
QY 929 GTCCCCGGAAGCCTGACCTTCTGACGAGCCTGAGAAATGACCTGCCCTGGCCAGC 988
DB 2793 TTCTTCCCTCTCTGCTGCTGCTGCTGCAATTAATGATTTGTTCTTCCCTATGGGATC 2852
QY CCTACTCTGTGTATAGATTAAGAGCGCTGCTGTGTGTGGAATAAAAAAAAAAAAAA 1048
DB 2853 CATCTGTTTGTATAACAATAAAGGCTCTGAGGAGATGTAAAAA 2912
QY 1049 AAAAAAAAAAAAAA 1060

DB 2913 AAAAAAAAAA 2924

RESULT 14

AAF32706

ID AAF32706 standard; cDNA; 2968 BP.

XX AAF32706;

AC

XX

DT 23-MAR-2001 (first entry)

XX

DE Human secreted protein gene 8 SEQ ID NO:18.

XX

Human; secreted protein; diagnosis; immunosuppressive; antiarthritic; anti-rheumatic; antiproliferative; cyostatic; cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial; virucide; fungicide; opthalmological; autoimmune disease; rheumatoid arthritis; hyperproliferative disorder; neoplasm; cardiovascular disorder; cardiac arrest; cerebrovascular disorder; cerebral ischaemia; infection; angiogenesis; nervous system disorder; Alzheimer's disease; skin aging; ocular disorder; corneal infection; wound healing; food additive; preservative; ss.

XX

OS Homo sapiens.

XX

PN MO200077255-A1.

XX

XX 21-DEC-2000.

PD

XX

PF 01-JUN-2000; 2000MO-US014926.

XX

PR 11-JUN-1999; 99US-0138628P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Ruben SM, Komatsoulis GA;

XX

DR WPI; 2001-025337/03.

XX

DR P-PSDB; AAB64429.

XX

XX

Isolated nucleic acid molecule encoding a human secreted protein is used in preventing, treating or ameliorating a medical condition.

XX

Claim 1; Page 477; 593pp: English.

XX

The polynucleotide sequences given in AAF32699 to AAF32747 encode the human secreted proteins given in AAB64422 to AAB64470. AAB64471 to AAB64548 represent human secreted polypeptide sequences and proteins homologous to them, which are given in the exemplification of the present invention. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Examples of activities include: antiarthritic; immunosuppressive; anti-rheumatic; antiproliferative; cyostatic; cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial; virucide; fungicide; and opthalmological. The polynucleotides and polypeptides can be used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. AAF32690 to AAF32698 and AAB64421 represent sequences used in the exemplification of the present invention

SQL Sequence 2968 BP; 679 A; 816 C; 797 G; 671 T; 0 U; 5 Other;

Query Match 4.7%; Score 49.6; DB 4; Length 2968;

Best Local Similarity 53.6%; Pred. No. 0.2;

Matches 103; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 869 GCCTGACACAGCCCTGAGAAATCCCGTAATATTATGACTTTGGGCTTGGCCCTGT 928

DB 2733 GCCCGGAGGAGCCCTCTTGCATAGGAGTAGATTGGCATTTAAAGCTCACCTTTT 2792

QY 929 GTCCCCGAGCGCTGTGACCTTCTGACGCGAGCTGAGAAATGACTGCCGCCACG 988

DB 2793 TTCTTCCCTCTCTGTGCTGTGCTGTGACATTAATGATTTGTTCTTCCCTATGGGATC 2852

QY 989 CCTACTGTGTGAATAGAAATTAAGCCGCGGTGTCTGTGSAAAAAAAAAAAAAA 1048

DB 2853 CATCTGTTTGTAAACAATTAAGCGCTGAGGAGCTTAAAAAAAAAAAAAAA 2912

QY 1049 AAAAAAAAAA 1060

DB 2913 AAAAAAAAAA 2924

RESULT 15

AA188830

ID AA188830 standard; cDNA; 375 BP.

XX

AC AA188830;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 8890.

XX

XX

Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorders; arthritis; inflammation; ss.

XX

OS Homo sapiens.

XX

PN WO200164835-A2.

XX

PD 07-SEP-2001.

XX

PF 26-FEB-2001; 2001WO-US004927.

XX

PR 28-FEB-2000; 2000US-00515126.

XX

PR 18-MAY-2000; 2000US-00577409.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Dzmanac RT;

XX

DR WPI; 2001-514838/56.

XX

DR P-PSDB; AAO08899.

XX

PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukemia, inflammation and immune disorders.

XX

PS Claim 1; SEQ ID NO 8890; 1399pp + Sequence Listing; English.

XX

The invention relates to human polynucleotides (AA179941-AA193841) and the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activity/inhibitin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format

CC directly from WIP0 at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 375 BP; 130 A; 87 C; 65 G; 65 T; 0 U; 28 Other;

Query Match 4.6%; Score 48.6; DB 4; Length 375;

Best Local Similarity 60.0%; Pred. No. 0.19; Mismatches 0; Gaps 0;

Matches 78; Conservative 0; Indels 52; Gaps 0;

QY 931 CCCCCGACGCTGTGACACCTCTGACGCGCTGAGAACTGACCTGCGCCCGAGCCC 990

Db 197 CCCCCTGAATCCCAAGCACTTTGGAGGCTGGCGAGAACTGCTGACCCAGTGCATAAC 256

QY 991 TACTCTGTGTAATAGAAATAAGGCGCTGCTGTCTGTGAAAAAAAAAAAAAAAA 1050

Db 257 CATCTGGGCAACATAGCAGACCATGTCTGTNNNNAAAAAAAAAAAAAAAAAAAA 316

QY 1051 AAAAAAAAAA 1060

Db 317 AAAAAAAAAA 326

Search completed: November 22, 2004, 08:24:08
Job time : 889 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 07:49:15 ; Search time 208 Seconds
(without alignments)
3622.289 Million cell updates/sec

Title: US-09-597-920B-1
Perfect score: 1060
Sequence: 1 gactctgccttcgaggggcc.....aaaaaaaaaaaaaaaa 1060

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA:*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	416.8	39.3	528	4	US-09-023-655-565
2	51.4	4.8	7218	1	US-08-232-463-14
3	43.8	4.1	1926	4	US-09-249-585A-4
4	43.8	4.1	1931	2	US-09-130-114-2
5	42.4	4.0	1503	4	US-09-907-794A-220
6	42.4	4.0	1503	4	US-09-905-125A-220
7	42.4	4.0	1503	4	US-09-902-775A-220
8	42.4	4.0	1503	4	US-09-906-700-220
9	42.4	4.0	1503	4	US-10-140-002-363
10	42.4	4.0	1503	4	US-09-903-603A-220
11	42	4.0	977	4	US-09-311-021-103
12	42	4.0	3083	4	US-09-155-246-1
13	42	4.0	4494	4	US-09-620-312D-861
14	41.8	3.9	1505	1	US-07-915-246-1
15	41.8	3.9	1600	1	US-07-861-458C-37
16	41.8	3.9	1607	6	5196333-3
17	41.6	3.9	1160	3	US-08-995-159-1
18	41.6	3.9	1160	3	US-09-545-605-1
19	41.2	3.9	966	2	US-08-921-382-1
20	41.2	3.9	966	4	US-09-386-380-1
21	41.2	3.9	3489	2	US-08-728-323A-1
22	41.2	3.9	3489	3	US-09-298-568-1
23	41.2	3.9	3489	4	US-09-410-399-1
24	41.2	3.9	3489	4	US-09-894-273-1
25	41.2	3.9	32207	2	US-08-770-379-20
26	41.2	3.9	32207	3	US-08-757-669A-20
27	41.2	3.9	32207	3	US-09-230-371A-20

28	41	3.9	1366	3	US-09-200-965-1	Sequence 1, Appli
29	40.8	3.8	3028	4	US-09-548-938A-2	Sequence 2, Appli
30	40.4	3.8	997	4	US-09-800-729-14	Sequence 14, Appli
31	40.4	3.8	1071	4	US-09-800-729-45	Sequence 45, Appli
32	40.4	3.8	1086	4	US-09-800-729-48	Sequence 48, Appli
33	40.4	3.8	1138	4	US-09-800-729-44	Sequence 44, Appli
34	40.4	3.8	1149	4	US-09-800-729-47	Sequence 47, Appli
35	40.4	3.8	1798	4	US-09-797-906-1	Sequence 1, Appli
36	40.4	3.8	7218	1	US-08-232-463-14	Sequence 14, Appli
37	40.4	3.8	16442	3	US-08-781-891-208	Sequence 208, App
38	40.4	3.8	16442	4	US-09-618-166-208	Sequence 208, App
39	40.2	3.8	1851	4	US-09-149-476-62	Sequence 62, Appli
40	40.2	3.8	2920	3	US-09-276-400-1	Sequence 1, Appli
41	40.2	3.8	2920	3	US-09-448-076-1	Sequence 1, Appli
42	40.2	3.8	2920	3	US-09-702-572-1	Sequence 1, Appli
43	40	3.8	2744	3	US-09-071-101-1	Sequence 1, Appli
44	40	3.8	2744	4	US-09-369-618-1	Sequence 1, Appli
45	40	3.8	2744	3	US-09-369-617-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-023-655-565
Sequence 565, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
TITLE OR INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESSES:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 565:
SEQUENCE CHARACTERISTICS:
LENGTH: 528 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: HIPNOTO1
CLONE: 240885
US-09-023-655-565
Query Match 39.3%; Score 416.8; DB 4; Length 528;

Best Local Similarity 99.3%; Pred. No. 8.9e-95;
Matches 429; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 599 CAGAGGCTCTCTGATGCGACGCGGAGTATGTAATGTCCTCCAGGAACTGATCTG 658
Db 1 CAGAGGCTCTCTGATGCGACGCGGAGTATGTAATGTCCTCCAGGAACTGATCTG 60
QY 659 GAGCGGCTTAAGACTGAGCTGCGCCCTGAGTCCAGAGGAGAGAGAGTGAAG 718
Db 61 GAGCGGCTTAAGACTGAGCTGCGCCCTGAGTCCAGAGGAGAGAGAGTGAAG 120
QY 719 AGGGGGCTCCAGATTAGAGAACTGCGAGAGCTGAACTGAGGCTGTGAGCCGAGT 778
Db 121 AGGGGGCTCCAGATTAGAGAACTGCGAGAGCTGAACTGAGGCTGTGAGCCGAGT 180
QY 779 CTGCTCGAGACGAGGCTGCTGAGG-ACGAGTGAAGCTGAGGAGTGAAGTGAAGT 837
Db 181 CTGCTCGAGACGAGGCTGCTGAGGAGAGCTGAGGAGTGAAGTGAAGTGAAGT 240
QY 838 GGTCTCAGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 897
Db 241 GGTCTCAGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 898 TAACTTATATCACTTTGGGCTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 957
Db 301 TAACTTATATCACTTTGGGCTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 958 CAGCGTGAATGAGCTGCGCCCTGAGCCGAGGCTGCTGCTGCTGCTGCTGCTG 1017
Db 361 CAGCGTGAATGAGCTGCGCCCTGAGCCGAGGCTGCTGCTGCTGCTGCTGCTG 420
QY 1018 CGTGTCTGTGTG 1029
Db 421 CGTGTCTGTGTG 432

RESULT 2
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300

TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 4.8%; Score 51.4; DB 1; Length 7218;
Best Local Similarity 7.1%; Pred. No. 0.0044;
Matches 22; Conservative 168; Mismatches 119; Indels 0; Gaps 0;

QY 41 TCCGAGCTCCCTGCGAGATGAGAGGCGCATCTGTCCTCCCTGCGGCTGCTG 100
Db 1032 TCCGAGCTTGGCTGCGAGGTGAGAGGAGCTTGCATATTTTTTTTTTTTTTTT 1091
QY 101 TGTGCGCCATCTGCGCATGTGATGAGCACTGTGTGCACTGCCAGAGCT 160
Db 1092 TT 1151
QY 161 CTAAGACAGACATCTCAGATAGTTGTATCAAGGAGGATCCAGTTCAAAGGCTC 220
Db 1152 TT 1211
QY 221 ACAAGTGGCCCTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 280
Db 1212 TT 1271
QY 281 AGCCAGACTGCTCCCATCCCAAGATCCCGAGCCCTTGGGGCTCCAGAGCG 340
Db 1272 TT 1331
QY 341 CATCTTCCC 349
Db 1332 TTTTTTTTTT 1340

RESULT 3
US-09-249-585A-4/C
Sequence 4, Application US/09249585A
Patent No. 6417002
GENERAL INFORMATION:
APPLICANT: Horlick, Robert
TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
FILE REFERENCE: 0867/0D905
CURRENT APPLICATION NUMBER: US/09/249,585A
CURRENT FILING DATE: 1999-02-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin version 3.0
SEQ ID NO 4
LENGTH: 1926
TYPE: DNA
ORGANISM: Epstein Barr Virus
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(1926)
OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4

Query Match 4.1%; Score 43.8; DB 4; Length 1926;
Best Local Similarity 47.0%; Pred. No. 0.2; Indels 152; Gaps 0;
Matches 135; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 552 GAGTCATGATGATTAAGTGAAGTTCGAGAGCGGAGAGAGAGAGAGAGAGAGAG 611
Db 575 GAGAGACGGAG 516
QY 612 GAGTGGACGCGGAGAGATGATGTGTCAGAGAACTGATCTGAGAGCGGCTAGAC 671

Db 515 GGACGAGGACGGGAGGACGGGAGGACGGGAGGACGAGACGGGAGGACGGGAG 456
Qy 672 TGAGCTGCGCCCTGATGTTCCAGGAGGACGAGGAGGAGGAGGAGGAGGAGGAG 731
Db 455 GGAGGACGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 396
Qy 732 TTACGAGATCTGACGAGAGCTGAACTGAGGAGGAGGAGGAGGAGGAGGAGGAG 791
Db 395 CGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 336
Qy 792 AGGCTTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 838
Db 335 GGACGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 289

RESULT 4

US-09-130-114-2/c
Sequence 2, Application US/09130114
Patent No. 5976807
GENERAL INFORMATION:
APPLICANT: Horlick, Robert A.
APPLICANT: Damaj, Bassem B.
APPLICANT: Robbins, Alan K.
TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
FILE REFERENCE: 0867/ID903US1
CURRENT APPLICATION NUMBER: US/09/130,114
CURRENT FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1931
TYPE: DNA
ORGANISM: EBNA
US-09-130-114-2

Query Match 4.1%; Score 43.8; DB 2; Length 1931;
Best Local Similarity 47.0%; Pred. No. 0.2;
Matches 135; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

Qy 552 GGAGTCATTTGATGATTAAGTGAACGTTCCGAGAGCGGGAGAGCGGAGGAGGAG 611
Db 575 GGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 516
Qy 612 GGATGCGACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 671
Db 515 GGACGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 456
Qy 672 TGAGCTGCGCCCTGATGTTCCAGGAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 731
Db 455 GGAGGACGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 396
Qy 732 TTACGAGATCTGACGAGAGCTGAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 791
Db 395 CGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 336
Qy 792 AGGCTTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 838
Db 335 GGACGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 289

RESULT 5

US-09-907-794A-220
Sequence 220, Application US/09907794A
Patent No. 6635468
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Baton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Flavaro, Ellen

APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/907,794A
CURRENT FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 220
LENGTH: 1503
TYPE: DNA
ORGANISM: Homo sapiens
US-09-907-794A-220

Query Match 4.0%; Score 42.4; DB 4; Length 1503;
Best Local Similarity 66.3%; Pred. No. 0.41;
Matches 61; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 969 TGACCTGCTGCGCCCGCCGCTACTCTGTGTAATGAATTAAGGCTGCTGTCTGT 1028
Db 1385 TGGCTGCGCTGGGAGCACTCTCTGTCGAGGAGCAATTAAGCCGCGGAGACT 1444
Qy 1029 GGAAG 1060

; PRIOR APPLICATION NUMBER: PCT/US00/04414
 ; PRIOR FILING DATE: 2000-02-22

APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William

```
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P33081C59
CURRENT APPLICATION NUMBER: US/10/140,002
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 363
LENGTH: 1503
TYPE: DNA
ORGANISM: Homo Sapien
US-10-140-002-363

Query Match
Best Local Similarity 4.0%; Score 42.4; DB 4; Length 1503;
Matches 61; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 969 TGACCTGCGCTGGCCCGCCCTACTCTGTGTAATAGATAAGCCCTGCTGTCTGT 1028
DB 1385 TGGCGCTGGCGCTGGGACACCTCTCTGCGCAGAGGCAATAAAGCCGCGGACCT 1444
QY 1029 GGAATTAAGGCTGCTGTCTGTGTAATAGATAAGCCCTGCTGTCTGT 1060
DB 1445 TGAATTAAGGCTGCTGTCTGTGTAATAGATAAGCCCTGCTGTCTGT 1476

RESULT 10
US-09-903-603A-220
Sequence 220, Application US/09903603A
Patent No. 6767995
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Macher, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: GNB.161822C12
CURRENT APPLICATION NUMBER: US/09/903, 603A
CURRENT FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143, 048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145, 698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146, 222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
```

```
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 220
LENGTH: 1503
TYPE: DNA
ORGANISM: Homo sapiens
US-09-903-603A-220

Query Match
Best Local Similarity 4.0%; Score 42.4; DB 4; Length 1503;
Matches 61; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 969 TGACCTGCGCTGGCCCGCCCTACTCTGTGTAATAGATAAGCCCTGCTGTCTGT 1028
DB 1385 TGGCGCTGGCGCTGGGACACCTCTCTGCGCAGAGGCAATAAAGCCGCGGACCT 1444
QY 1029 GGAATTAAGGCTGCTGTCTGTGTAATAGATAAGCCCTGCTGTCTGT 1060
DB 1445 TGAATTAAGGCTGCTGTCTGTGTAATAGATAAGCCCTGCTGTCTGT 1476

RESULT 11
US-09-311-021-103
Sequence 103, Application US/09311021
Patent No. 6706869
GENERAL INFORMATION:
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Fechtel, Kim
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: GI 6300-11A
CURRENT APPLICATION NUMBER: US/09/311, 021
CURRENT FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 268
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 103
LENGTH: 977
TYPE: DNA
ORGANISM: Homo sapiens
US-09-311-021-103

Query Match
Best Local Similarity 4.0%; Score 42; DB 4; Length 977;
Matches 48; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1003 TGAATTAAGGCTGCTGTCTGTGTAATAGATAAGCCCTGCTGTCTGT 1060
DB 919 TGAATTAAGGCTGCTGTCTGTGTAATAGATAAGCCCTGCTGTCTGT 976
```



```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/915,246
FILING DATE: 19920716
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pochopien, Donald J.
REGISTRATION NUMBER: 32,167
REFERENCE/DOCKET NUMBER: 91 P 1125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312 707-8889
TELEFAX: 312 707-9155
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1505 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Brassica napus
STRAIN: Westar
DEVELOPMENTAL STAGE: Somatic
TISSUE TYPE: Root
US-07-915-246-1
```

```

Query Match      3.9% Score 41.8; DB 1; Length 1505;
Best Local Similarity 47.2%; Pred. No. 0.58;
Matches 127; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
```

```

QY 552 GAGATCCATTGATGATTAAGTGAACGTTCCGAGAGAGCGGAGAGCGCAGAAAGCGTCTT 611
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 900 GAGAGAGCTGAGAGACATGTTGAGAGTGAGAGAGAGCGAGAGAGTGTTGAGAGAGCT 959
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 612 GATGAGAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 671
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 960 GCGCGTCCGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 672 TGAGCTGCGCGCCCTGAGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 731
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1020 GAGAGTGGCGGTGAGTGCATGAGATGAGAGTGTGAGAGCGGTGAGAAATGAGCGGTGA 1079
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 732 TTACGAGATCTGAGAGAGAGTGAATGAGAGCGCTGTGAGAGCGCGAGTCTGTCTGAACC 791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1080 GAGAGTGAAGTGCACACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1139
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 792 AGGCTTGCCTGGAGCGGCTGAGCTGGGCA 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1140 GATATGAGAGTGGCGGTGAGAGTGGACA 1168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 15
US-07-861-458C-37

Sequence 37, Application US/07861458C
Patent No. 6232061

GENERAL INFORMATION:

APPLICANT: Marchionni, Mark Andrew

APPLICANT: Johnson, Carl D.

TITLE OF INVENTION: HOMOLOGY CLONING

NUMBER OF SEQUENCES: 142

CORRESPONDENCE ADDRESS:

STREET: 225 Franklin Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 50Z or 55SX

OPERATING SYSTEM: MS-DOS (Version 5.0)

SOFTWARE: WordPerfect (Version 5.1)

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/861,458C
FILING DATE: 04/01/92
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 04585/014001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-07-861-458C-37
```

```

Query Match      3.9% Score 41.8; DB 3; Length 1600;
Best Local Similarity 71.4%; Pred. No. 0.6;
Matches 55; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
```

```

QY 984 CCAGCCCTACTGTGTATAGATTAAGCGCTGCGTGTCTGTGGAAGAAAAA 1043
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1522 CCATGCAATGTTGTGTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1044 AAAAAAAAAAAAAAAAAA 1060
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1582 AAAAAAAAAAAAAAAAAA 1598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: November 22, 2004, 12:08:20
Job time : 209 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 10:12:50 ; Search time 616 Seconds
(without alignments)
9298.443 Million cell updates/sec

Title: US-09-597-920B-1

Perfect score: 1060

Sequence: 1 gacctgcgccttgaggggcc.....aaaaaaaaaaaaaaaaaaaa 1060

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3627888 seqs, 2701811610 residues

Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	949.2	89.5	1224	9	US-09-739-907-13
2	949.2	89.5	1224	11	US-09-938-671-13
3	424	40.0	457	15	US-10-102-524-1371
4	416.8	39.3	528	16	US-10-641-643-565
5	271.6	25.6	319	9	US-09-796-692-5333
6	271.6	25.6	319	14	US-10-040-862-5333
7	271.6	25.6	319	15	US-10-057-475B-5333
8	271.6	25.6	319	16	US-10-154-884B-5333
9	271.6	25.6	319	17	US-10-764-324-5333
C 10	208	19.6	421	9	US-09-796-692-5582
C 11	208	19.6	421	14	US-10-040-862-5582
C 12	208	19.6	421	16	US-10-057-475B-5582

C 13	208	19.6	421	16	US-10-154-884B-5582	Sequence 5582, Ap
C 14	208	19.6	421	17	US-10-764-324-5582	Sequence 5582, Ap
C 15	205.2	19.4	400	16	US-10-242-535A-9083	Sequence 9083, Ap
C 16	205.2	19.4	400	16	US-10-085-783A-9083	Sequence 9083, Ap
C 17	117	11.0	117	9	US-09-796-692-5736	Sequence 5736, Ap
C 18	117	11.0	117	14	US-10-040-862-5736	Sequence 5736, Ap
C 19	117	11.0	117	16	US-10-057-475B-5736	Sequence 5736, Ap
C 20	117	11.0	117	16	US-10-154-884B-5736	Sequence 5736, Ap
C 21	117	11.0	117	17	US-10-764-324-5736	Sequence 5736, Ap
C 22	49.6	4.7	2968	9	US-09-925-301-584	Sequence 584, App
C 23	49.6	4.7	2968	15	US-10-106-698-1500	Sequence 1500, Ap
C 24	48.8	4.6	500	18	US-10-425-115-83126	Sequence 83126, A
C 25	47.4	4.5	1419	17	US-10-437-963-25990	Sequence 25990, A
C 26	47.4	4.5	3196	9	US-09-782-980-50	Sequence 50, App1
C 27	47.4	4.5	3196	17	US-10-806-018-50	Sequence 50, App1
C 28	46	4.3	477	10	US-09-918-995-10907	Sequence 10907, A
C 29	45.8	4.3	2022	17	US-10-437-963-21276	Sequence 21276, A
C 30	45.6	4.3	468	10	US-09-918-995-21830	Sequence 21830, A
C 31	45	4.2	496	16	US-10-240-425-154	Sequence 154, App
C 32	45	4.2	777	14	US-10-184-644-148	Sequence 148, App
C 33	45	4.2	777	14	US-10-184-634-148	Sequence 148, App
C 34	44.8	4.2	173	9	US-09-834-975-589	Sequence 589, App
C 35	44.8	4.2	189	9	US-09-834-975-390	Sequence 390, App
C 36	44.8	4.2	848	9	US-09-823-245A-186	Sequence 186, App
C 37	44.6	4.2	424	9	US-09-867-701-1117	Sequence 1117, Ap
C 38	44.6	4.2	472	10	US-09-918-995-16588	Sequence 16588, A
C 39	44.4	4.2	910	14	US-10-198-846-6483	Sequence 6483, A
C 40	44.2	4.2	178	17	US-10-437-963-23303	Sequence 23303, A
C 41	44.2	4.2	512	18	US-10-425-115-131750	Sequence 131750, App1
C 42	44.2	4.2	1338	14	US-10-097-065-19	Sequence 19, App1
C 43	44.2	4.2	1338	15	US-10-372-876-19	Sequence 19, App1
C 44	44.2	4.2	1537	10	US-09-820-788-1	Sequence 1, App1
C 45	44.2	4.2	1677	16	US-10-264-237-401	Sequence 401, App

ALIGNMENTS

RESULT 1
US-09-739-907-13
; Sequence 13, Application US/09739907
; Patent No. US20010012869A1
GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/739,907
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1205)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-739-907-13
Query Match 89.5%, Score 949.2; DB 9; Length 1224;
Best Local Similarity 98.9%; Pred. No. 1.5e-261;
Matches 1039; Conservative 0; Mismatches 4; Indels 8; Gaps 8;

OY	1	AACTGCGCCCTTGAAGGGGCTTAAGGGGATGACGACGCTGCTCCAGACTCCCTCGAATG	60
Db	181	GACTCTGCGCTTGAAGGGGCTTAAGGGGATGACGACGCTGCTCCAGACTCCCTCGAATG	240
OY	61	GAGAGGCGCATCTGTGCTCCCTGCGTGTGGGGCTCTGTGCTGTGCCATCTGGCCATG	120
Db	241	GAGAGGCGCATCTGTGCTCCCTGCGTGTGGGGCTCTGTGCTGTGCCATCTGGCCATG	300
OY	121	TTGATGGGACGTGTGTGTGACATGGCCACAGACTGGCCAGGGCTCTTAAGACAGACATCTCA	180
Db	301	TTGATGGGACGTGTGTGTGACATGGCCACAGACTGGCCAGGGCTCTTAAGACAGACATCTCA	360
OY	181	GATAGTTTGTATTC-AAGGGGATCCAGATTGAACGGGCTGACACGGTGTGCGCCCTGGCC	239
Db	361	GATAGTTTGTATTCGAAGGGGATCCAGATTGAACGGGCTGACACGGTGTGCGCCCTGGCC	420
OY	240	ACCTGCGCTAACCGACCTGTACCTCTTACCACCCCTTGAGCCAGCCAGACCTGTCCCAT	299
Db	421	ACCTGCGCTAACCGACCTGTACCTCTTACCACCCCTTGAGCCAGCCAGACCTGTCCCAT	480
OY	300	CCCAAGATCCCCGGGCGAGCCCTTTGGGGGGTCCACACGGGACATTTCCCGGGGGATTC	359
Db	481	CCCAAGATCCCCGGGCGAGCCCTTTGGGGGGTCCACACGGGACATTTCCCGGGGGATTC	540
OY	360	TGATGCTCCAACTGTGTGTGCGAGCTTACGAGAACGAGAACCAAGCTTGAAGATGACGA	419
Db	541	TGATGCTCCAACTGTGTGTGCGAGCTTACGAGAACGAGAACCAAGCTTGAAGATGACGA	599
OY	420	TGAGGATGAGAGCGCATTCACAAACCCAGGCTTACTGTGTGCTTCTTGACAGACCCC	479
Db	600	TGA-GATAGGACGATCATACAAACCCAGGCTTACTGTGTGCTTCTTGACAGACCCC	658
OY	480	GGCCACTGACCTGTGCGCCCATCAGCTCTGTGACTGTGACACCCCTGTGCATCCGACAG	539
Db	659	GGCCACTGACCTGTGCGCCCATCAGCTCTGTGACTGTGACACCCCTGTGCATCCGACAG	717
OY	540	TGCTTTCTCCATGAGTCTTATGATTAAGTAAAGTTCCGAGAGACGGGGAGAGCGC	599
Db	718	TGCTTTCTCCATGAGTCTTATGATTAAGTAAAGTTCCGAGAGACGGGGAGAGCGC	777
OY	600	AGAAAGCTCTCTGATGTGGCAGCGGGAGATATGTGAATGTGCCAGAACTGCATCTCTGG	659
Db	778	AGAAAGCTCTCTGATGTGGCAGCGGGAGATATGTGAATGTGCCAGAACTGCATCTCTGG	837
OY	660	AGCGGCTAAGACTGAGCTGTGCGCCCTTGAATTTCCAGAGGCGAGAGAAATGAGAGAA	719
Db	838	AGCGGCTAAGACTGAGCTGTGCGCCCTTGAATTTCCAGAGGCGAGAGAAATGAGAGAA	897
OY	720	GGGGGCTCCAGATTACGAAGATCTGACAGAGCTGAACTGAGGGCCTGTGAGAGCCGAACTC	779
Db	898	GGGGGCTCCAGATTACGAAGATCTGACAGAGCTGAACTGAGGGCCTGTGAGAGCCGAACTC	957
OY	780	TGTCCTGGAACCAAGGCTTGTGCTGTGGACGAGCTGAGCTGTGGGAGTGAAGTGTGCGG	839
Db	958	TGTCCTGGAACCAAGGCTTGTGCTGTGGACGAGCTGAGCTGTGGGAGTGAAGTGTGCGG	1017
OY	840	TCCTTACATGGGCTCTGTGCTTGTCTTCAAGCTTGAACACAGCTTGAAGAAATCCCCGCTA	899
Db	1018	TCCTTACATGGGCTCTGTGCTTGTCTTCAAGCTTGAACACAGCTTGAAGAAATCCCCGCTA	1077
OY	900	ACTTATTTACATTGGGGGTTGCGGCTGTGTCCCCCGAAGCTCTGACCTTGTGAGCGCA	959
Db	1078	ACTTATTTACATTGGGGGTTGCGGCTGTGTCCCCCGAAGCTCTGAGCTTGTGAGCGCA	1135
OY	960	GCTTGAGATGACCTGTGCGCTGTGGCCCGACCCCTTACTTGTGTATTAAGATTAAGGCTTGC	1019
Db	1136	GCTTGAGATGACCTGTGCGCTGTGGCCCGACCCCTTACTTGTGTATTAAGATTAAGGCTTGC	1193
OY	1020	TGTGCTGTGAAAAAATTTTTTTTTTTTTTTT 1050	
Db	1194	TGTGCTGTGTAAAAAATTTTTTTTTTTTTTTT 1224	

```

RESULT2
US-09-938-671-13
; Sequence 13, Application US/09938671
; Publication No. US20040002066A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/938, 671
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,592
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1205)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-938-671-13

```

Query Match	89.5%	Score 949.2	DB 11	Length 1224
-------------	-------	-------------	-------	-------------

Best Local Similarity: 58.5%; EREC: NO: 1.00 201;
Matches 1039; Conservative: 0; Mismatches: 4; Indels: 8; Gaps: 8;

QY	1	GACTGTGCGCTTGAAGGGGCTTAAGGGGTTGACGACAGCCGTGCTCCGAGCTCCCTCGAGATG	60
Db	181	GACTGTGCGCTTGAAGGGGCTTAAGGGGTTGACGACAGCCGTGCTCCGAGCTCCCTCGAGATG	240
QY	61	GAGAGAGCCATCTTGTCCTCCCTGCGCTGGGGGCTCTGTGCTTGCCTATCTGGCCATG	120
Db	241	GAGAGAGCCATCTTGTCCTCCCTGCGCTGGGGGCTCTGTGCTTGCCTATCTGGCCATG	300
QY	121	TTGATGACACTGTGTGTGTCACCTGCCACAGACTGCCAGGCTCTCTACGACGACATCTCTA	180
Db	301	TTGATGACACTGTGTGTGTCACCTGCCACAGACTGCCAGGCTCTCTACGACGACATCTCTA	360
QY	181	GATAGTTTGATTC-AAAGGGGATTCAGATTCAAAAGGCTTCACAGGTTGGCCCCCTGGCC	239
Db	361	GATAGTTTGATTC-AAAGGGGATTCAGATTCAAAAGGCTTCACAGGTTGGCCCCCTGGCC	420
QY	240	AACGTGCCATCCACCTGTACCTCTTACCCACCCCTGAAGCAGCAGACTGTCTCCCAT	299
Db	421	AACGTGCCATCCACCTGTACCTCTTACCCACCCCTGAAGCAGCAGACTGTCTCCCAT	480
QY	300	CCCAAGATTCGCCGAGCCCTTGTGGGGGCTCCCAACCGACAGCCATCTTCCGGCGGGATTTC	359
Db	481	CCCAAGATTCGCCGAGCCCTTGTGGGGGCTCCCAACCGACAGCCATCTTCCGGCGGGATTTC	540
QY	360	TGATGTGCCAACAGTGTGGCGAGCTACGAGAACGAGAACCCAGCTGTAGAGATGCAGA	419
Db	541	TGATGTGCCAACAGTGTGGCGAGCTACGAGAACGAGAACCCAGCTGTAGAGATGCAGA	599
QY	420	TGAGGATGAGACGACTATCACAAACCCAGGCTACTGGTGTGTCTTCTGTACAGACACCCC	479
Db	600	TGA-GATAGGACGACTATCACAAACCCAGGCTACTGGTGTGTCTTCTGTACAGACACCCC	658
QY	480	GGCCACTGACACTGTGCCCCATGAGCTCTGACATCCAGACACCCCTGGGATCCGAGACAG	539
Db	659	GGCCACTGACACTGTGCCCCATGAGCTCTGACATCCAGACACCCCTGGGATCCGAGACAG	717

Qy 540 TGCCTTCATGAGTCCATGATGATTAACGTGACGCTCCGAGAGCGGAGAGCGC 599
Db 718 TGCCTTCATGAGTCCATGATGATTAACGTGACGCTCCGAGAGCGGAGAGCGC 777
Qy 600 AGAAGCGTCTCGATGAGCAGCGGAGATGATGATGTCCTCCAGAACTGTCATCTCG 659
Db 778 AGAAGCGTCTCGATGAGCAGCGGAGATGATGATGTCCTCCAGAACTGTCATCTCG 837
Qy 660 AGCGGCTAAGCTGAGCGCTGCGGCTGATGTTCCAGAGAGCGAGAGAGTGAAGAGA 719
Db 838 AGCGGCTAAGCTGAGCGCTGCGGCTGATGTTCCAGAGAGCGAGAGTGAAGAGA 897
Qy 720 GGGGCTCCAGATTAAGAGAACTGCAAGAGCTGAATGAGGCTCTGAGAGCGGAGATC 779
Db 898 GGGGCTCCAGATTAAGAGAACTGCAAGAGCTGAATGAGGCTCTGAGAGCGGAGATC 957
Qy 780 TGTCTGGAACGAGGCTTGCCTGAGAGCGCTGAGAGCTGAGAGTGAAGTGTGAGG 839
Db 958 TGTCTGGAACGAGGCTTGCCTGAGAGCGCTGAGAGCTGAGAGTGAAGTGTGAGG 1017
Qy 840 TCCCTCAATGAGCGCTGCGGCTGCTGAGCTGACCAAGCGCTGAAGAAATCCCGGCTA 899
Db 1018 TCCCTCAATGAGCGCTGCGGCTGCTGAGCTGACCAAGCGCTGAAGAAATCCCGGCTA 1077
Qy 900 ACTTATTAATCACTTGGGGGTTGCGGCTGTGTCCCGAAGCTCTGACCTTCTGACCA 959
Db 1078 ACTTATTAATCACTTGGGGGTTGCGGCTGTGTCCCGAAGCTCTGACCTTCTGACCA 1135
Qy 960 GCTGAGAGATGACCTGCGGCTGCGGCTGCTGATGATGATGATGATGATGATGATG 1019
Db 1136 GCTGAGAGATGACCTGCGGCTGCGGCTGCTGATGATGATGATGATGATGATGATG 1193
Qy 1020 TGTCTGCTGAG 1050
Db 1194 TGTCTGCTGAG 1224

RESULT 3
US-10-102-524-1371
; Sequence 1371, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Mannion, Paul A.
; APPLICANT: Algate, Paul A.
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102.524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1371
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 449_
; OTHER INFORMATION: n = A,T,C or G
US-10-102-524-1371

Query Match 40.0%; Score 424; DB 15; Length 457;
Best Local Similarity 99.8%; Pred. No. 3.6e-11;
Matches 424; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 397 GAACACGCTTGTGAGATGAGATGAGATGAGACGCTATCACACCCGAGGCTACTG 456
Db 33 GAACACGCTTGTGAGATGAGATGAGATGAGACGCTATCACACCCGAGGCTACTG 92
Qy 457 GTGCTGCTTCTGACAGACCCGCGCACTAGCACTGTGCCCATCAGCTCTGCACTC 516

Db 93 GTGGTGCTTCTGACAGACCCCGGCACTAGCACTGTGCCCCATCAGCTCTGCACTC 152
Qy 517 AGCACCCCTGGCATCCAGAGACAGTCTCTTCTCATGAGTCCATGATGATTAAGTGAAC 576
Db 153 AGCACCCCTGGCATCCAGAGACAGTCTCTTCTCATGAGTCCATGATGATTAAGTGAAC 212
Qy 577 GTTCCGAGAGCGGGGAGAGCGGCAAGAGTCTGATGAGAGCGGAGAGTGAAT 636
Db 213 GTTCCGAGAGCGGGGAGAGCGGCAAGAGTCTGATGAGAGCGGAGAGTGAAT 272
Qy 637 GTGTCCAGAACTGATCTCTGAGCGGCTAAGACTGAGCTGCGGCTGATGTTCCAG 696
Db 273 GTGTCCAGAACTGATCTCTGAGCGGCTAAGACTGAGCTGCGGCTGATGTTCCAG 332
Qy 697 GAGCAGAGAGAGTGAAGAGAGAGGCGGCTCCAGATTGAGAACTTCAGAGAGCTGAAC 756
Db 333 GAGCAGAGAGAGTGAAGAGAGAGGCGGCTCCAGATTGAGAACTTCAGAGAGCTGAAC 392
Qy 757 TGAGGCGCTGTGAGAGCGGAGATCTGTCTGAAACAGGCTTGCGGAGCGGCTGAGCTG 816
Db 393 TGAGGCGCTGTGAGAGCGGAGATCTGTCTGAAACAGGCTTGCGGAGCGGCTGAGCTG 452
Qy 817 GGCAG 821
Db 453 GGCAG 457

RESULT 4
US-10-641-643-565
; Sequence 565, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/641.643
; FILING DATE: 14-Aug-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 565:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 528 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HIPNOTO1

CLONE: 240885
SEQUENCE DESCRIPTION: SEQ ID NO: 565 :
US-10-641-643-565

Query Match 39.3%; Score 416.8; DB 16; Length 528;
Best Local Similarity 99.3%; Pred. No. 4,4e-109;
Matches 429; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 599 CAGAACGCTCTGATGAGCGCGGAGATGTAATGTGTCCCGAGAACTGCATCTG 658
DB 1 CAGAACGCTCTGATGAGCGCGGAGATGTAATGTGTCCCGAGAACTGCATCTG 60
QY 659 GACCGGCTTAAGACTGAGCTGCGCCCTGAGTTCCAGAGAGAGAGAAAGTGAAG 718
DB 61 GACCGGCTTAAGACTGAGCTGCGCCCTGAGTTCCAGAGAGAGAGAAAGTGAAG 120
QY 719 AGGGGGCTCCAGATTACGAAATCTGAGAGAGCTGAACCTAGGGCCCTGTGAGAGCCGAGT 778
DB 121 AGGGGGCTCCAGATTACGAAATCTGAGAGAGCTGAACCTAGGGCCCTGTGAGAGCCGAGT 180
QY 779 CTGTCTGAGAACGAGCTGCTGCGG-ACGGCTGAGCTGGGAGAGTGAAGTGCCTG 837
DB 181 CTGTCTGAGAACGAGCTGCTGCGGAGAGCTGAGAGCTGAGAGTGAAGTGCCTG 240
QY 838 GGTCTTCAATGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 897
DB 241 GGTCTTCAATGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 898 TAACCTATTATCACTTTGGGGTTGGGCTGTGTCCCGGAAAGCTGTGACCTTTGACG 957
DB 301 TAACCTATTATCACTTTGGGGTTGGGCTGTGTGTCCCGGAAAGCTGTGACCTTTGACG 360
QY 958 CAGCCTGAGAAATGAGCTGCGCTGAGCGCCGAGCCCTACTCTGTGTATGATTAAGGCTG 1017
DB 361 CAGCCTGAGAAATGAGCTGCGCTGAGCGCCGAGCCCTACTCTGTGTATGATTAAGGCTG 420
QY 1018 CGTGTGTCTGTG 1029
DB 421 CGTGTGTCTGTG 432

RESULT 5
US-09-796-692-5333
Sequence 5333, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077, 001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903

PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5333
LENGTH: 319
TYPE: DNA
ORGANISM: Homo sapiens
US-09-796-692-5333

Query Match 25.6%; Score 271.6; DB 9; Length 319;
Best Local Similarity 98.6%; Pred. No. 1.6e-67;
Matches 274; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 544 TTCTCATGAGAGTCCATTGATGATTACGTGAACCTTCCGAGAGCGCGGAGAGCCGAGAA 603
DB 42 TCCACAGTGGAGTCCATTGATGATTACGTGAACCTTCCGAGAGCGCGGAGAGCCGAGAA 101
QY 604 GCGTCTGAGATGAGCGCGGAGATGATGATGTGCTCCAGAGACTGCATCTGAGCG 663
DB 102 GCGTCTGAGATGAGCGCGGAGATGATGATGTGCTCCAGAGACTGCATCTGAGCG 161
QY 664 GCTAAGACTGAGCTGCGCCCTGAGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 723
DB 162 GCTAAGACTGAGCTGCGCCCTGAGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 221
QY 724 GCTCCAGATTACGAAATCTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 783
DB 222 GCTCCAGATTACGAAATCTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 281
QY 784 CTGGAACGAGGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 821
DB 282 CTGGAACGAGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 319

RESULT 6
US-10-040-862-5333
Sequence 5333, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903

;; PRIOR APPLICATION NUMBER: US 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: US 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; PRIOR APPLICATION NUMBER: US 09/796,692
;; PRIOR FILING DATE: 2001-03-01
;; NUMBER OF SEQ ID NOS: 10467
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO: 5333
;; LENGTH: 319
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-040-862-5333

Query Match 25.6%; Score 271.6; DB 14; Length 319;
Best Local Similarity 98.6%; Pred. No. 1.6e-67;
Matches 274; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 544 TTCTCATGAGTGCATTGATGATTACGTAACGTTCCGAGAGCGGGAGAGCCAGAA 603
Db 42 TCCACAGTGGAGTCCATTGATGATTACGTAACGTTCCGAGAGCGGGAGAGCCAGAA 101
Qy 604 GCGTCTGATGAGCGAGCGGGAGTATGTGTCTCCAGAACTGCATCTGAGCG 663
Db 102 GCGTCTGATGAGCGAGCGGGAGTATGTGTCTCCAGAACTGCATCTGAGCG 161
Qy 664 GCTAAGACTGAGCTGCGCCCTGAGTTCCTCCAGAGCGAGAACTGAGAGAGGGG 723
Db 162 GCTAAGACTGAGCTGCGCCCTGAGTTCCTCCAGAGCGAGAACTGAGAGAGGGG 221
Qy 724 GCTCAGATTGAGAACTGCGAGAGCTGAACCTGAGAGCGCGAGTCTGTC 783
Db 222 GCTCAGATTGAGAACTGCGAGAGCTGAACCTGAGAGCGCGAGTCTGTC 281
Qy 784 CTGGAACCAAGGCTTGCCTGGAGCGGCTGAGCTGGGCG 821
Db 282 CTGGAACCAAGGCTTGCCTGGAGCGGCTGAGCTGGGCG 319

RESULT 7
US-10-057-475B-5333
;; Sequence 5333, Application US/10057475B
;; Publication No. US20040002068A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; APPLICANT: Clapper, Jonathan David
;; APPLICANT: Wang, Aijun
;; APPLICANT: Ordenez, Nadia
;; APPLICANT: Carter, Lauren
;; APPLICANT: McNeill, Patricia Dianne
;; TITLE OF INVENTION: Corixa Corporation
;; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;; FILE REFERENCE: 014058-014402US
;; CURRENT APPLICATION NUMBER: US/10/057,475B
;; PRIOR FILING DATE: 2002-01-22
;; PRIOR APPLICATION NUMBER: US 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: US 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: US 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: US 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: US 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: US 60/206,201

;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: US 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 10979
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO: 5333
;; LENGTH: 319
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-057-475B-5333

Query Match 25.6%; Score 271.6; DB 16; Length 319;
Best Local Similarity 98.6%; Pred. No. 1.6e-67;
Matches 274; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 544 TTCTCATGAGTGCATTGATGATTACGTAACGTTCCGAGAGCGGGAGAGCCAGAA 603
Db 42 TCCACAGTGGAGTCCATTGATGATTACGTAACGTTCCGAGAGCGGGAGAGCCAGAA 101
Qy 604 GCGTCTGATGAGCGAGCGGGAGTATGTGTCTCCAGAACTGCATCTGAGCG 663
Db 102 GCGTCTGATGAGCGAGCGGGAGTATGTGTCTCCAGAACTGCATCTGAGCG 161
Qy 664 GCTAAGACTGAGCTGCGCCCTGAGTTCCTCCAGAGCGAGAACTGAGAGAGGGG 723
Db 162 GCTAAGACTGAGCTGCGCCCTGAGTTCCTCCAGAGCGAGAACTGAGAGAGGGG 221
Qy 724 GCTCAGATTGAGAACTGCGAGAGCTGAACCTGAGAGCGCGAGTCTGTC 783
Db 222 GCTCAGATTGAGAACTGCGAGAGCTGAACCTGAGAGCGCGAGTCTGTC 281
Qy 784 CTGGAACCAAGGCTTGCCTGGAGCGGCTGAGCTGGGCG 821
Db 282 CTGGAACCAAGGCTTGCCTGGAGCGGCTGAGCTGGGCG 319

RESULT 8
US-10-154-884B-5333
;; Sequence 5333, Application US/10154884B
;; Publication No. US2004000561A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; APPLICANT: Retter, Marc W.
;; APPLICANT: Corixa Corporation
;; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;; FILE REFERENCE: 014058-013521US
;; CURRENT APPLICATION NUMBER: US/10/154,884B
;; PRIOR FILING DATE: 2002-05-23
;; PRIOR APPLICATION NUMBER: US 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: US 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: US 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: US 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: US 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: US 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: US 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 60/222,903

;; PRIOR FILING DATE: 2000-08-03
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 11290
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 5333
;; LENGTH: 319
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-154-984B-5333

Query Match 25.6%; Score 271.6; DB 16; Length 319;
Best Local Similarity 98.6%; Pred. No. 1.6e-67;
Matches 274; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 544 TTCTCCATGGAGCTCATTGATGATTAAGTGAACCTTCGAGAGCGGGAGCGCAGAA 603
DB 42 TCACAGATGGAGTTCATTGATGATTAAGTGAACCTTCGAGAGCGGGAGCGCAGAA 101
QY 604 GCGTCTGTGATGGACCGCGGAGTATGTGATGTCTCCAGAACTGCATCTTGAGCG 663
DB 102 GCGTCTGTGATGGACCGCGGAGTATGTGATGTCTCCAGAACTGCATCTTGAGCG 161
QY 664 GCTTAAGCTGAGCTGCGCCGCTGAGTTCCAGAGGCAAGAGAGAGAGAGAGG 723
DB 162 GCTTAAGCTGAGCTGCGCCGCTGAGTTCCAGAGGCAAGAGAGAGAGAGAGG 221
QY 724 GCTTCAAGATTACAGAACTGTCAGAGAGCTGAAGAGGCGCTGTGAGAGCCGAGTCTGTC 783
DB 222 GCTTCAAGATTACAGAACTGTCAGAGAGCTGAAGAGGCGCTGTGAGAGCCGAGTCTGTC 281
QY 784 CTGGAACCAAGGCTTGCTGGGACGCGCTGAGCTGGGCGAG 821
DB 282 CTGGAACCAAGGCTTGCTGGGACGCGCTGAGCTGGGCGAG 319

RESULT 9
US-10-764-324-5333
; Sequence 5333, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0

;; SEQ ID NO 5333
;; LENGTH: 319
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-764-324-5333

Query Match 25.6%; Score 271.6; DB 17; Length 319;
Best Local Similarity 98.6%; Pred. No. 1.6e-67;
Matches 274; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 544 TTCTCCATGGAGCTCATTGATGATTAAGTGAACCTTCGAGAGCGGGAGCGCAGAA 603
DB 42 TCACAGATGGAGTTCATTGATGATTAAGTGAACCTTCGAGAGCGGGAGCGCAGAA 101
QY 604 GCGTCTGTGATGGACCGCGGAGTATGTGATGTCTCCAGAACTGCATCTTGAGCG 663
DB 102 GCGTCTGTGATGGACCGCGGAGTATGTGATGTCTCCAGAACTGCATCTTGAGCG 161
QY 664 GCTTAAGCTGAGCTGCGCCGCTGAGTTCCAGAGGCAAGAGAGAGAGAGG 723
DB 162 GCTTAAGCTGAGCTGCGCCGCTGAGTTCCAGAGGCAAGAGAGAGAGAGG 221
QY 724 GCTTCAAGATTACAGAACTGTCAGAGAGCTGAAGAGGCGCTGTGAGAGCCGAGTCTGTC 783
DB 222 GCTTCAAGATTACAGAACTGTCAGAGAGCTGAAGAGGCGCTGTGAGAGCCGAGTCTGTC 281
QY 784 CTGGAACCAAGGCTTGCTGGGACGCGCTGAGCTGGGCGAG 821
DB 282 CTGGAACCAAGGCTTGCTGGGACGCGCTGAGCTGGGCGAG 319

RESULT 10
US-09-796-692-5582/c
; Sequence 5582, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077,001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5582
; LENGTH: 421
; TYPE: DNA


```
QY 822 CTGGAAGTGGCTCTGGGGTCTCTACATAGCGCTCTGCCCTTGTCTCCAGCTGACACAGC 881
DB 421 CTGGAAGTGGCTCTGGGGTCTCTACATAGCGCTCTGCCCTTGTCTCCAGCTGACACAGC 362
QY 882 CTGGAAGTGGCTCTGGGGTCTCTACATAGCGCTCTGCCCTTGTCTCCAGCTGACACAGC 941
DB 361 CTGGAAGTGGCTCTGGGGTCTCTACATAGCGCTCTGCCCTTGTCTCCAGCTGACACAGC 302
QY 942 TCTGACCTTCTGACGCGAGCTGAGAAATGACCTGCGCCGCGAGCCTACTCTGTGTA 1001
DB 301 TCTGACCTTCTGACGCGAGCTGAGAAATGACCTGCGCCGCGAGCCTACTCTGTGTA 242
QY 1002 ATGAAATTAAGGCTGCTGTGTCTGTG 1029
DB 241 ATGAAATTAAGGCTGCTGTGTCTGTG 214
```

RESULT 13

```
US-10-154-884B-5582/C
; Sequence 5582, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Paul A.
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154, 884B
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5582
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-5582
```

```
Query Match 19.6%; Score 208; DB 16; Length 421;
Best Local Similarity 100.0%; Pred. No. 2.8e-49;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 822 CTGGAAGTGGCTCTGGGGTCTCTACATAGCGCTCTGCCCTTGTCTCCAGCTGACACAGC 881
DB 421 CTGGAAGTGGCTCTGGGGTCTCTACATAGCGCTCTGCCCTTGTCTCCAGCTGACACAGC 362
QY 882 CTGGAAGTGGCTCTGGGGTCTCTACATAGCGCTCTGCCCTTGTCTCCAGCTGACACAGC 941
DB 361 CTGGAAGTGGCTCTGGGGTCTCTACATAGCGCTCTGCCCTTGTCTCCAGCTGACACAGC 302
```

```
QY 942 TCTGACCTTCTGACGCGAGCTGAGAAATGACCTGCGCCGCGAGCCTACTCTGTGTA 1001
DB 301 TCTGACCTTCTGACGCGAGCTGAGAAATGACCTGCGCCGCGAGCCTACTCTGTGTA 242
QY 1002 ATGAAATTAAGGCTGCTGTGTCTGTG 1029
DB 241 ATGAAATTAAGGCTGCTGTGTCTGTG 214
```

RESULT 14

```
US-10-764-324-5582/C
; Sequence 5582, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Paul A.
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; PRIOR FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5582
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-764-324-5582
```

```
Query Match 19.6%; Score 208; DB 17; Length 421;
Best Local Similarity 100.0%; Pred. No. 2.8e-49;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 822 CTGGAAGTGGCTCTGGGGTCTCTACATAGCGCTCTGCCCTTGTCTCCAGCTGACACAGC 881
DB 421 CTGGAAGTGGCTCTGGGGTCTCTACATAGCGCTCTGCCCTTGTCTCCAGCTGACACAGC 362
QY 882 CTGGAAGTGGCTCTGGGGTCTCTACATAGCGCTCTGCCCTTGTCTCCAGCTGACACAGC 941
DB 361 CTGGAAGTGGCTCTGGGGTCTCTACATAGCGCTCTGCCCTTGTCTCCAGCTGACACAGC 302
QY 942 TCTGACCTTCTGACGCGAGCTGAGAAATGACCTGCGCCGCGAGCCTACTCTGTGTA 1001
DB 301 TCTGACCTTCTGACGCGAGCTGAGAAATGACCTGCGCCGCGAGCCTACTCTGTGTA 242
QY 1002 ATGAAATTAAGGCTGCTGTGTCTGTG 1029
DB 241 ATGAAATTAAGGCTGCTGTGTCTGTG 214
```

RESULT 15

US-10-242-535A-9083
; Sequence 9083, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9083
; LENGTH: 400
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (378)..(378)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-9083

Query Match 19.4%; Score 205.2; DB 16; Length 400;
Best Local Similarity 97.0%; Pred.No.1.8e-48;

Matches 230; Conservative 0; Mismatches 4; Indels 3; Gaps 2;

Qy	613	GATGCGAGCCGGAGATATGTATGTATGTCACAGAACTGATCTGAGCGGCTAAGACT	672
Db	164	GATGCGAGCCGGAGATATGTATGTATGTCACAGAACTGATCTGAGCGGCTAAGACT	223
Qy	673	GAGCTGCGCCCTGATGTTCCAGAGCAGAGAAAGTGAAGAGAGGGGCTCCAGAT	732
Db	224	GAGCTGCGCCCTGATGTTCCAGAGCAGAGAAAGTGAAGAGAGGGGCTCCAGAT	283
Qy	733	TACGAGAAATCTGAGAGAGCTGAATGAGGGGCTGTGAGGGCCGAGTCTCTGGAACCA	792
Db	284	TACGAGAAATCTGAGAGAGCTGAATGAGGGGCTGTGAGGGCCGAGTCTCTGGAACCA	343
Qy	793	GGCTTGCTGAGAC-GGCTGAGCTGAGGCACTGAA--GTGCTCTGAGGCTCTCAC	846
Db	344	GGCTTGCTGAGACGGGCTGAGCTGAGGCACTGAAAGTGGCTCTGGGCTCTCAC	400

Search completed: November 22, 2004, 13:40:22
Job time : 617 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 07:46:40 ; Search time 6715 Seconds
(without alignments)
5752.214 Million cell updates/sec

Title: US-09-597-920B-1

Perfect score: 1060

Sequence: 1 gactctgccccttaggggccc.....aaaaaaaaaaaaaaaaaaaa 1060

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 1821985908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST: *
1: gb_ests1: *
2: gb_ests2: *
3: gb_hnc: *
4: gb_ests3: *
5: gb_ests4: *
6: gb_ests5: *
7: gb_ests6: *
8: gb_gsa1: *
9: gb_gsa2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	1027.4	96.9	1534	3	CR624409 full-length
2	1027.4	96.9	1564	3	CR597721 full-length
3	1027.4	96.9	1604	3	CR592470 full-length
4	968.4	91.4	1207	3	CR602091 full-length
5	814.8	76.9	1090	4	BM810074 AGENCOURT
6	666	62.8	702	9	AY403398 Homo sapi
7	645.2	60.9	896	1	AL558417
8	633.6	58.8	983	1	AL558417
9	632.4	59.7	635	4	BG751757 602730483
10	625.6	59.0	937	1	AL580592
11	625.2	59.0	922	1	AL580240
12	612.2	57.8	651	2	AM612429
13	612	57.7	878	1	AL558866
14	606.4	57.2	610	4	BM715456
15	578	54.5	811	5	BK37371
16	575.6	54.3	916	5	BK447994
17	575	54.2	775	5	BI753306
18	572	54.0	846	5	BK437370
19	549	51.8	771	4	BI838499
20	547	51.6	924	4	BI910865
21	543.2	51.2	1096	4	BM562211
22	534.4	50.4	544	6	CD370683
23	525.6	49.6	1080	5	BQ072651
24	524.4	49.5	884	5	BX369483

25	519.4	49.0	817	4	BI911404
26	519.2	48.0	702	9	AY403399
27	516.8	48.8	874	4	BI909045
28	512	48.3	630	2	BF906497
29	509	48.0	517	4	BM673935
30	493.6	46.6	929	4	BC386066
31	477.6	45.1	540	5	BO581351
32	477.2	45.0	826	4	BI838054
33	467.4	44.1	643	2	AM384820
34	466	44.0	665	2	AM384832
35	459.8	43.4	568	2	AM970405
36	454	42.8	518	1	AI922013
37	445	42.0	898	5	BK400523
38	444.4	41.9	816	4	BI911769
39	440	41.5	456	1	AM002632
40	433.2	40.9	649	6	CA426247
41	432.2	40.8	477	2	AM296085
42	429.6	40.5	450	2	BF508034
43	423.8	40.0	657	4	BI909736
44	420.8	39.7	827	4	BC397816
45	418	39.4	892	5	BX330901

ALIGNMENTS

RESULT 1	CR624409	1534 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	CR624409				
DEFINITION	full-length cDNA clone CSOCAP06YF09 of Thymus of Homo sapiens (human)				
ACCESSION	CR624409				
VERSION	CR624409.1	GI:50505216			
KEYWORDS	HTC; CNSLT cDNA.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jeessee, J., and Polayes, D.				
TITLE	1 (bases 1 to 1534)				
JOURNAL	Full-length cDNA libraries and normalization				
REMARK	Unpublished				
	Contact : Feng Liang Email : fliang@life.techn.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paraday Avenue				
	2 (bases 1 to 1534)				
REFERENCE	Genoscope.				
AUTHORS	Genoscope.				
TITLE	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage ;				
JOURNAL	BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
COMMENT	- Web : www.genoscope.cns.fr				
	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime				
	end enriched, double-strand cDNA was digested with Not I and cloned				
	into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library				
	was normalized. Library was constructed by Life Technologies, a				
	division of Invitrogen.				
FEATURES	Location/Qualifiers				
source	1. 1534				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CSOCAP06YF09"				
	/tissue_type="Thymus"				
	/plasmid="pCMVSPORT_6"				
ORIGIN					
Query Match	96.9%; Score 1027.4; DB 3; Length 1534;				
Best Local Similarity	99.9%; Pred. No. 4.2e-229;				
Matches 1028; Conservative	0; Mismatches 1; Indels 0; Gaps 0;				
1	GACTCTGCCCTTAGGGGCTAGGGGTGCAGCCGCTCCGAGCCCTCCGAGATG 60				
234	GACTCTGCCCTTAGGGGCTAGGGGTGCAGCCGCTCCGAGCTCCGAGATG 293				

QY 901 CTTATTATCACTTTGGGGTTGCGCTGTGTCTCCCGAAGCTCTGACACTTCTGACGAG 960
DB 1140 CTTATTATCACTTTGGGGTTGCGCTGTGTCTCCCGAAGCTCTGACACTTCTGACGAG 1199
QY 961 CTTGAGAAATGACCTGCGCTGCGCCCGAAGCTCTGCTGTGTATTAATGAATAAGGCTGCGT 1020
DB 1200 CTTGAGAAATGACCTGCGCTGCGCCCGAAGCTCTGCTGTGTATTAATGAATAAGGCTGCGT 1259
QY 1021 GTGTCTGTG 1029
DB 1260 GTGTCTGTG 1268

RESULT 4
CR602091 1207 bp mRNA linear HTC 21-JUL-2004
DEFINITION Full-length cDNA clone CS0DJ005YC09 of T cells (Jurkat cell line)
ACCESSION CR602091
VERSION CR602091.1 GI:50482898
KEYWORDS HTC; CNSLT cDNA
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1207)
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL: <http://fulllength.invitrogen.com/InvitrogenCorporation1600>
Faraday Avenue
2 (bases 1 to 1207)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(AT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source location/Qualifiers
1..1207
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ005YC09"
/issue_type="T cells (Jurkat cell line) cot
10-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 91.4%; Score 968.4; DB 3; Length 1207;
Best Local Similarity 99.9%; Pred. No. 2.4e-215;
Matches 969; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 241 CTTGCTACCCACCTGTGCACTCTCTACCCACCCCTGAGCGAGCAAGCTGTCCCATC 300
DB 478 CTTGCTACCCACCTGTGCACTCTCTACCCACCCCTGAGCGAGCAAGCTGTCCCATC 537
QY 301 CCAAGATCCCGGAGCCCTTGGGGGCTCCACCGAGCGCATCTTCCGGGGATTC 360
DB 538 CCAAGATCCCGGAGCCCTTGGGGGCTCCACCGAGCGCATCTTCCGGGGATTC 597
QY 361 GATGTGCTCAACAGTGTGGGAGCTACGAGAAAGAGGAACACCTGTGAGATGCAAT 420
DB 598 GATGTGCTCAACAGTGTGGGAGCTACGAGAAAGAGGAACACCTGTGAGATGCGAT 657
QY 421 GAGATGAGAGCACTATTCACAAACCCAGGCTACTGTGTGCTTCCCTGACAGACCCG 480
DB 658 GAGATGAGAGCACTATTCACAAACCCAGGCTACTGTGTGCTTCCCTGACAGACCCG 717
QY 481 GCCACTAGCACTGTGCGCCCATGACTCTGTGCACTGACACCCCTGGCATCCGACAGT 540
DB 718 GCCACTAGCACTGTGCGCCCATGACTCTGTGCACTGACACCCCTGGCATCCGACAGT 777
QY 541 GCCCTTCCATGAGTGCATTTGATGATTAAGTAAAGTTCGGAGAGCGGGAGAGCGCA 600
DB 778 GCCCTTCCATGAGTGCATTTGATGATTAAGTAAAGTTCGGAGAGCGGGAGAGCGCA 837
QY 601 GAAGCGTCTGTGATGAGCGAGCGGAGATATGTGAATGTGCCAGAACTGCATCTCTGA 660
DB 838 GAAGCGTCTGTGATGAGCGAGCGGAGATATGTGAATGTGCCAGAACTGCATCTCTGA 897
QY 661 GCGGCTTAAGACTGAGCGCTGCGCCCTGAGTTCCAGAGGCGAAGATGAGAAAG 720
DB 898 GCGGCTTAAGACTGAGCGCTGCGCCCTGAGTTCCAGAGGCGAAGATGAGAAAG 957
QY 721 GGGGCTCCAAATTAACGAATCTGCAAGAGCTGAACCTGAGGGCTGTGAGGCGAGTCT 780
DB 958 GGGGCTCCAAATTAACGAATCTGCAAGAGCTGAACCTGAGGGCTGTGAGGCGAGTCT 1017
QY 781 GTCTGGAACCAAGCTTGTGCTGAGAGCGCTGAGCTGAGAGAGCTGTGAGGCGAGTCT 840
DB 1018 GTCTGGAACCAAGCTTGTGCTGAGAGCGCTGAGCTGAGAGAGCTGTGAGGCGAGTCT 1077
QY 841 CTTGACATGCGCTGTGCTGCTTGTGCTGACGCTGACCAACAGCTGAGAAATCCCGCTAA 900
DB 1078 CTTGACATGCGCTGTGCTGCTTGTGCTGACGCTGACCAACAGCTGAGAAATCCCGCTAA 1137
QY 901 CTTATTATCACTTTGGGGTTGCGCTGTGTCTCCCGAAGCTCTGACACTTCTGACGAG 960
DB 1138 CTTATTATCACTTTGGGGTTGCGCTGTGTCTCCCGAAGCTCTGACACTTCTGACGAG 1197
QY 961 CTTGAGAAATG 970
DB 1198 CTTGAGAAATG 1207

RESULT 5
BM810074 1090 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6579672 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5470018
DEFINITION 5', mRNA sequence.
ACCESSION BM810074
VERSION BM810074.1 GI:19126897
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1090)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gsabds-f@mail.nih.gov
Tissue Procurement: DCTD/DTF

[illegible]

LOCUS	AL558417	896 bp	mRNA	linear	EST 02-APR-2004
DEFINITION	AL558417 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED				
ACCESSION	AL558417				
VERSION	AL558417.3				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1. (bases 1 to 896)				
JOURNAL	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
COMMENT	Full-length cDNA libraries and normalization				
	Unpublished (2001)				
	On Feb 15, 2001 this sequence version replaced gi:31280216.				
	Contact: Genoscope				
	Genoscope - Centre National de Sequencage				
	BP 191 91006 EVRY cedex - France				
	Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr				
	1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five primam				
	end enriched, double-strand cDNA was digested with Not I and clone				
	into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library				
	was normalized. Library was constructed by life Technologies, a				
	division of Invitrogen. This sequence belongs to sequence cluster				
	8019.f				

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?c=CS0DDJ0055B05QPI&c=8019.f>.

FEATURES	Location/Qualifiers
SOURCE	1..896
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="CS0D005YC09"
	/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
	/cell_line="JURKAT"
	/clone_id="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
	/note="1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN	
Query Match	60.9%; Score 645.2; DB 1; Length 896;
Beat Local Similarity	99.2%; Pred. No. 7.6e-140;
Matches 655; Conservative	3; Mismatches 1; Indels 1; Gaps 1;
Db	1 GACTCTGCGCTTTGAGGGGGCTTAGGGGTGCACGCCAGCTGCTCCGAGCTCCTTCGACATG 60
Db	238 GACTCTGCGCTTGAGGGGGCTTAGGGGTGCACGCCAGCTGCTCCGAGCTCCTTCGACATG 297
Qy	61 GAGAGGGGCATCTGTCGTCCCTGGTGTGCTGGGGTCTCTGTCGTGCGCATCTGGCCATG 120
Db	298 GAGAGGGGCATCTGTCGTCCCTGGTGTGCTGGGGTCTCTGTCGTGCGCATCTGGCCATG 357
Qy	121 TTGATGTGCATCTGTGTGTGCATCTGCGCACAGATCGCAGGCTCTTCAAGACATCTCTCA 180
Db	358 TTGATGTGCATCTGTGTGTGCATCTGCGCACAGATCGCAGGCTCTTCAAGACATCTCTCA 417
Qy	191 GATAATTGTATCCAAGGGGGCATCAAGTTCAAAGCGGCTTACAACAGTTGCCCTTGCGCA 240
Db	418 GATAGTTGTATCCAAGGGGGCATCAAGTTCAAAGCGGCTTACAACAGTTGCCCTTGCGCA 477
Qy	241 CCTGCTTACCACCTCTGTACCTCTCTTACCACCCCTTGAGCGAGCGAAGCTGTCCTCCATC 300
Db	478 CCTGCTTACCACCTCTGTACCTCTCTTACCACCCCTTGAGCGAGCGAAGCTGTCCTCCATC 537
Qy	301 CCAAGATCCCCGCGAGCCCCTTGAGGGGGTCCCGACCGAGAGCCATCTTCCCGCGGGATTCT 360
Db	538 CCAAGATCCCCGCGAGCCCCTTGAGGGGGTCCCGACCGAGAGCCATCTTCCCGCGGGATTCT 597
Qy	361 GATGTGTCMAACAGTGTGGCGAGCTAGAGAACAGAAACAGGCTGTGAGATGACAGT 420
Db	598 GATGTGTCMAACAGTGTGGCGAGCTAGAGAACAGAAACAGGCTGTGAGATGACAGT 657
Qy	421 GAGATGAGACGACTATCAACACCCAGGACTACTGTGTGTGTTCTCTGACAGACACCCCG 480
Db	658 GAGATGAGACGACTATCAACACCCAGGACTACTGTGTGTGTTCTCTGACAGACACCCCG 717
Qy	481 GCCACTACACCTGCTGCGCCCATCAGCTCTGSCACTCAGACACCCCTGGGATCCGAGACAGT 540
Db	718 GCCACTACACCTGCTGCGCCCATCAGCTCTGSCACTCAGACACCCCTGGGATCCGAGACAGT 777
Qy	541 GCCTTCTTCATGAAGTCCATTGATGATTACGTGAACGTTCCGGAAGAGGGGAGAGCGCA 600
Db	778 GCCTTCTTCATGAAGTCCATTGATGATTACGTGAACGTTCCGGAAGAGGGGAGAGCGCA 837
Qy	601 GAAGCGTCTTGATGATGAGCGAGCCGGAGATATGTAATGTGTCCAGAACTGCATCTCTGA 660
Db	838 GAAGCGTCTTGATGATGAGCGAGCCGGAGATATGTAATGTGT-CCMRGACATGCATCTCTGA 896
RESULT 8	
LOCUS	BX447995 983 bp mRNA linear EST 06-MAY-2004
DEFINITION	BX447995 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
ACCESSION	BX447995
VERSION	BX447995.2 GI:47064222
KEYWORDS	EST.

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (baes 1 to 983)
AUTHORS	Ll.W.B., Gruber,C., Jeesee,J. and Polayes,D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	On May 22, 2003 this sequence version replaced gi:31027882. Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqef@genoscope.cns.fr Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, division of Invitrogen. This sequence belongs to sequence cluster 8019.f For more information about this cluster, see http://www.genoscope.cns.fr/cdna?c=CS0DM005CH07QPlsc=8019.f .
FEATURES	Location/Qualifiers
source	1..983 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DM005YP13" /issue_type="FETAL LIVER" /dev_stage="fetal" /clone_lib="Homo sapiens FETAL LIVER" /note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
ORIGIN	
Query Match	59.8%; Score 633.6; DB 5; Length 983;
Best Local Similarity	98.6%; Pred. No.3,9e-137;
Matches	655; Conservative 4; Mismatches 3; Indels 2; Gaps 2
Dy	1 GACTCTGCGCCCTTGAGGGGCGTTAGGGGTCAGGCACGCTGTCCAGACTCCCTGCAGATG 60
Dd	239 GACTCTGCCCTTGAGGGGCGTTAGGGGTCAGGCACGCTGTCCAGACTCCCTGCAGATG 298
Dy	61 GAGGAGGCGCATCTCGTGTCCCTCGCTGCTGGGGCTTCGTGCTGCTGCCATCCTGGCCATG 120
Dd	299 GAGGAGGCGCATCTCGTGTCCCTCGCTGCTGGGGCTTCGTGCTGCTGCCATCCTGGCCATG 358
Dy	121 TTGATGCGCACTGTGTGTGTCACCTGCGACAGA CTGCGCAGGCTCTTAGAGACGACATCTCTCA 180
Dd	359 TTGATGCGCACTGTGTGTGTCACCTGCGACAGA CTGCGCAGGCTCTTAGAGACGACATCTCTCA 418
Dy	181 GATAGTTTGATCCAAAGGGGCGATTCAGTTCAAACGGGCTTCACACGGCTTGCCTCCCTGGACA 240
Dd	419 GATAGTTTGATCCAAAGGGGCGATTCAGTTCAAACGGGCTTCACACGGTTCGSCCTTGGCCA 478
Dy	241 CCTGCGCTAACCCACCTGTGACCTCTTCAACCCACCCCTGAGCCAGCCAGACCTGTCTCCCATC 300
Dd	479 CCTGCGCTAACCCACCTGTGACCTCTTCAACCCACCCCTGAGCCAGCCAGACCTGTCTCCCATC 538
Dy	301 CCAAGATCCCCCGCAGGCCCTTGTGGGGGCTTCCACCGCAGCGCCATCTTCCCGGGGGATTCT 360
Dd	539 CCAAGATCCCCCGCAGGCCCTTGTGGGGGCTTCCACCGCAGCGCCATCTTCCCGGGGGATTCT 598
Dy	361 GATGTGCGCAACAGTGTGGCGAGCTTAGAAGAA CGAGGAACCAAGCCTGTGAGGATGCAGAT 420
Dd	599 GATGTGCGCAACAGTGTGGCGAGCTTAGAAGAA CGAGGAACCAAGCCTGTGAGGATGCAGAT 658
Dy	421 GAGATGAGGAGCGACTATCAACAACCAGAGCTACCGGTGGGTCTTCTGAGACGACACCCCG 480
Dd	659 GAGATGAGGAGCGACTATCAACAACCAGAGCTACCGGTGGGTCTTCTGAGACGACACCCCG 718

QY	481	GGCACTACGACATGCGTCCCCCATCTCCTGTGACACTCAGACACCCCTGGGACATCCGAGACACT	540
Db	719	GCACATGACACACTGCTGCCCCCATCACTCTCTGTGACACTCAGACACCCCTGGGACATCCGAGACACT	778
QY	541	GCCTTCTTCATGAGAGTCCATTGATGATTAACGTAAAGTTCGCGAGAGCGGAGAGACGCA	600
Db	779	CCCTTCTTCATGAGATCCATTGATTAACGTAAAGTTCGCGAGAGCGGAGAGACGCA	837
QY	601	GAAGCGTCTTCGATGCGACGCCGGAGATGTGAATGTGTCCAGAGACTGCATCTGGA	660
Db	838	GAAGCGTCTTCGATGCGACGCCGGAGATGTGAATGTGTCCAGAGACTGCATCTGGA	896
QY	661	CGCG 664	
Db	897	CGCG 900	
RESULT 9			
LOCUS	BG751757		
DEFINITION	BG751757	635 bp	mRNA linear EST 15-MAY-2000
ACCESSION	602730483p1 NIH_MGC_43	635 bp	Homosapiens cDNA clone IMAGE:4874108 5',
VERSION	BG751757		mRNA sequence.
KEYWORDS	BG751757.1	GI:14062410	
SOURCE	EST.		
ORGANISM	Homosapiens (human)		
REFERENCE	Homosapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/.		
COMMENT	1 (bases 1 to 635)		
	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Ling Hong/Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		
	http://image.lnl.gov		
	Plate: LUCM1753	row: e	column: 21
	High quality sequence stop: 635.		
FEATURES	Location/Qualifiers		
source	1..635		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4874108"		
	/tissue_type="normal pigmented retinal epithelium"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_lib="NIH_MGC_43"		
	/note="Organ: eye; Vector: pOTB7, Site_1: XhoI, Site_2:		
	EcoRI; cDNA made by oligo-dT priming. Directionally		
	cloned into EcoRI/XhoI sites using the following 5'		
	adaptor: GGCACGAG(G). Library constructed by Ling Hong		
	in the laboratory of Gerald M. Rubin (University of		
	California, Berkeley) using ZAP-cDNA synthesis kit		
	(Stratagene) and Superscript II RT (Life Technologies).		
	Note: this is a NIH_MGC Library. "		
ORIGIN			
Query Match	59.7%	Score 632.4	DB 4; Length 635;
Best Local Similarity	99.8%	Pred. No. 7.1e-137;	
Matches 633; Conservative 0;	Mismatch 1;	Indels 0;	Gaps 0;
QY	126	GGCAGCTGTGTGAGACTGCGACAGACTGCGAGGCTCTCTAGACAGACATCTCTGATAG	185
Db	2	GGCAGCTGTGTGAGACTGCGACAGACTGCGAGGCTCTCTAGACAGACATCTCTGATAG	61
QY	186	TTGTGATTCAGAGGGGACATCCAGTTCAACGGCTTCACAGGTTGGCCCTTGACCACTTCG	245


```

Db      62 TTTGTATCCAAAGGGGCGATCCAGTTCAAAAGGCTTCACAGGTTGCCCCCTGGCCACCTGC 121
Qy      246 CTACCCACCTGTGACCTCTTACACCCCTTGAGGCGACGACTGCTCCCATCCAG 305
Db      122 CTACCCACCTGTGACCTCTTACACCCCTTGAGGCGACGACTGCTCCCATCCAG 181
Qy      306 ATCCCGCAGCCCTTGGGGGCTCCACCGGACGCCATCTTCCCGGGGGAATTGTGATG 365
Db      182 ATCCCGCAGCCCTTGGGGGCTCCACCGGACGCCATCTTCCCGGGGGAATTGTGATG 241
Qy      366 TGGCAACAGTGTGGCGAGCTACGAAACGAGAAACGAGCTGTAGATGTAGATGAGA 425
Db      242 TGGCAACAGTGTGGCGAGCTACGAAACGAGAAACGAGCTGTAGATGTAGATGAGA 301
Qy      426 TGAAGAGAGTATACAAACCCAGGCTACTGTGTGTGTTCTGACAGACCCCGGCGAC 485
Db      302 TGAAGAGAGTATACAAACCCAGGCTACTGTGTGTGTTCTGACAGACCCCGGCGAC 361
Qy      486 TAGCAGCTGTGCCCCCATCAGCTCTCTGACCTCAGACCCCTTGACAGACAGTGCCTT 545
Db      362 TAGCAGCTGTGCCCCCATCAGCTCTCTGACCTCAGACCCCTTGACAGACAGTGCCTT 421
Qy      546 CTCCATGAGTCCATTGATGATTAAGTGAACTTCCGAGAGCGGGGAGAGCGAGAAC 605
Db      422 CTCCATGAGTCCATTGATGATTAAGTGAACTTCCGAGAGCGGGGAGAGCGAGAAC 481
Qy      606 GTCTCTGATGAGGCGCGGGAGTATGTGAATGTGTCCAGGAACTGCATCCTGAGGGG 665
Db      482 GTCTCTGATGAGGCGCGGGAGTATGTGAATGTGTCCAGGAACTGCATCCTGAGGGG 541
Qy      666 TAAAGCTGAGGCTCCCGCCCTGAGTTCACAGAGGCAAGAGAGTGAAGAGAGGGG 725
Db      542 TAAAGCTGAGGCTCCCGCCCTGAGTTCACAGAGGCAAGAGAGTGAAGAGAGGGG 601
Qy      726 TCCAGATTACGAAATCTGCAGAGCTGAAGTGA 759
Db      602 TCCAGATTACGAAATCTGCAGAGCTGAAGTGA 635

```

```

RESULT 10
LOCUS    AL580592/c 937 bp mRNA linear EST 07-APR-2004
DEFINITION Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
            Homo sapiens cDNA clone CS0DJ015YE16 3-PRIME, mRNA sequence.
ACCESSION AL580592
VERSION   AL580592.1 GI:12946762
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 937)
AUTHORS  Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 Evry cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by life technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            8019.F
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdnas=CS0DJ015BC08NP1&c=8019.f.
            Location/Qualifiers
                +1..937
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DJ015YE16"

```

```

/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

```

Query Match 59.0%; Score 625.6; DB 1; Length 937;
Best Local Similarity 98.3%; Pred. No. 2.9e-135;
Matches 625; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy      394 GAGGAACACAGCTGTGAGATGACAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 453
Db      937 GAGGAACACAGCTGTGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 878
Qy      454 CTGATGTGCTTCTCTGACAGACCCCGGCGACATGACATGCTGCGCCATGAGCTCTGCA 513
Db      877 CTGATGTGCTTCTCTGACAGACCCCGGCGACATGACATGCTGCGCCATGAGCTCTGCA 818
Qy      514 CTACAGACCCCTTGACATCCGAGACAGTGCCTTCTCCATGAGATGATGATTAAGT 573
Db      817 CTACAGACCCCTTGACATCCGAGACAGTGCCTTCTCCATGAGATGATGATTAAGT 758
Qy      574 AACGTTCCGAGAGCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 633
Db      757 AACGTTCCGAGAGCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 698
Qy      634 AATGTGTCCAGAACTGCATCTGTGAGCGGCTGAAGCTGAGCTGCGCGCTGAGTTCC 693
Db      697 AATGTGTCCAGAACTGCATCTGTGAGCGGCTGAAGCTGAGCTGCGCGCTGAGTTCC 638
Qy      694 CAGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 753
Db      637 CAGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 578
Qy      754 AACTGAGGCGCTGTGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 813
Db      577 AACTGAGGCGCTGTGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 518
Qy      814 CTGGGAGCTGGAAGTGGCTCTGGGGTCTTACATGAGCGTCTGCTGCTGCAAGCTG 873
Db      517 CTGGGAGCTGGAAGTGGCTCTGGGGTCTTACATGAGCGTCTGCTGCTGCAAGCTG 458
Qy      874 ACAACAGCTGGAAGATCCCGGCTGAATTAATTAATTAATTAATTAATTAATTAAT 933
Db      457 ACAACAGCTGGAAGATCCCGGCTGAATTAATTAATTAATTAATTAATTAATTAAT 398
Qy      934 CCGAAGCTGTGACCTTGTGAGCGAGCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAG 993
Db      397 CCGAAGCTGTGACCTTGTGAGCGAGCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAG 338
Qy      994 TCTGTGTAATAGATTAAGGCTGCGTGTGTGTG 1029
Db      337 TCTGTGTAATAGATTAAGGCTGCGTGTGTGTG 302

```

```

RESULT 11
LOCUS    AL580240/c 922 bp mRNA linear EST 07-APR-2004
DEFINITION Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
            Homo sapiens cDNA clone CS0DJ005YC09 3-PRIME, mRNA sequence.
ACCESSION AL580240
VERSION   AL580240.3 GI:46258981
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 922)
AUTHORS  Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.

```


RESULT 14
LOCUS BM715456 610 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-EJ0-ahi-g-12-0-UI.r2 UI-E-EJ0 Homo sapiens cDNA clone
BM715456
ACCESSION BM715456
VERSION BM715456.1 GI:19028714
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 610)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resegen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1. 610
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-ahi-g-12-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_1lb="UI-E-EJ0"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-EJ0 is a subcloned cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
AGATTCACAGA; lens, CGATTACGA; eye anterior segment,
AATGCCCAT; optic nerve, CCATTAGTG; retina, CCGCG; Retina
foveal and macular, GTCC; RPE and choroid, ACCCTA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI)."

ORIGIN
Query Match 57.2%; Score 606.4; DB 4; Length 610;
Best Local Similarity 99.5%; Pred. No. 8.3e-131;
Matches 607; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 416 CAGATGAGATGAGACGATTCACAAACCCAGGCTACCTGGTGCTCTTCGACACACA 475
DB 1 CGGATGAGATGAGACGATTCACAAACCCAGGCTACCTGGTGCTCTTCGACACACA 60

QY 476 CCCCCGACCTAGACATGCTGCTGCCCATCAGCTCTGACACTAGACACCCCTGGCATCCGAG 535
DB 61 CCCCCGACCTAGACATGCTGCTGCCCATCAGCTCTGACACTAGACACCCCTGGCATCCGAG 120
QY 536 ACAGTGCCTTCTCCATGAGAGTCCATTTGATGATTAAGTGAACGTTCCGAGACGGGGAGA 595
DB 121 ACAGTGCCTTCTCCATGAGAGTCCATTTGATGATTAAGTGAACGTTCCGAGACGGGGAGA 180
QY 596 GCGCAGAACGCTCTTGATGAGACGCGGAGATATGTAATGTCTCCAGCAATCTGCATC 655
DB 181 GCGCAGAACGCTCTTGATGAGACGCGGAGATATGTAATGTCTCCAGCAATCTGCATC 240
QY 656 CTGAGCGGCTTAAGCTAGAGCTGCGGCTGAGTCTCCAGAGGACAGAGAAAGTGAAG 715
DB 241 CTGAGCGGCTTAAGCTAGAGCTGCGGCTGAGTCTCCAGAGGACAGAGAAAGTGAAG 300
QY 716 AAGAGGGGCTCCAGATTAAGAGAAATCTGACAGAGCTGAAGGCTGTGAAGGCG 775
DB 301 AAGAGGGGCTCCAGATTAAGAGAAATCTGACAGAGCTGAAGGCTGTGAAGGCG 360
QY 776 AGCTGTCTCTGAGAACAGAGCTTGGTGGAGGCTGAGCTGGGACGCTGAAAGTGGCT 835
DB 361 AGCTGTCTCTGAGAACAGAGCTTGGTGGAGGCTGAGCTGGGACGCTGAAAGTGGCT 420
QY 836 GGGGTCTCAGATGCGCTGCGCTTGGCTCAGCTGACCAACAGCCGAGAAATCCGCC 895
DB 421 GGGGTCTCAGATGCGCTGCGCTTGGCTCAGCTGACCAACAGCCGAGAAATCCGCC 480
QY 896 CGTAACTATTATCACTTTGGGGTTGCGGCTGTGCTCCGCAACGCTGCACTTCTGA 955
DB 481 CGTAACTATTATCACTTTGGGGTTGCGGCTGTGCTCCGCAACGCTGCACTTCTGA 540
QY 956 CGCAGCTGAGAAATGACCTGCTGCGGCGCCAGCCCTACTGTGTATGATTAAGAGCC 1015
DB 541 CGCAGCTGAGAAATGACCTGCTGCGGCGCCAGCCCTACTGTGTATGATTAAGAGCC 600
QY 1016 TGGCTGTGTC 1025
DB 601 TGGCTGTGTC 610

RESULT 15
LOCUS BX437371 811 bp mRNA linear EST 04-MAY-2004
DEFINITION BX437371 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP006YF09
5-PRIME, mRNA sequence.
ACCESSION BX437371
VERSION BX437371.2 GI:47003573
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 811)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
unpublished (2001)
COMMENT On May 15, 2003 this sequence version replaced gi:30777562.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 8019.f.
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna7s-CS0CAP006C05Q1fC=8019.f.
location/Qualifiers
1. 811
/organism="Homo sapiens"

FEATURES
source

```
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="CS0CAP006YF09"  
/tissue_type="THYMUS"  
/clone_lib="Homo sapiens THYMUS"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."
```

ORIGIN

```
Query Match      54.5%; Score 578; DB 5; Length 811;  
Best Local Similarity 100.0%; Pred. No. 3,7e-124;  
Matches 578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY      1 GACTCTGCGCCCTTGAGGGGCTTAGGGGTGCGACGCCAGCTTCCGAGCTCCCTGCAAGT 60  
      |||  
Db      234 GACTCTGCGCCCTTGAGGGGCTTAGGGGTGCGACGCCAGCTGCTCCGAGCTCCCTGCAAGT 293  
  
QY      61 GAGAGGGCCATCTGTGCTCCCTGCGTGTGGGGCTCTGTGCTGCTGCCATCCTGGGCATG 120  
      |||  
Db      294 GAGAGGGCCATCTGTGCTCCCTGCGTGTGGGGCTCTGTGCTGCTGCCATCCTGGGCATG 353  
  
QY      121 TTGATGGCATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
      |||  
Db      354 TTGATGGCATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 413  
  
QY      181 GATAGTTGTATCCAAAGGGGATCCAGTTCAAAGGGCTCAAGGCTGCGGCTGCGGCA 240  
      |||  
Db      414 GATAGTTGTATCCAAAGGGGATCCAGTTCAAAGGGCTCAAGGCTGCGGCTGCGGCA 473  
  
QY      241 CCTGCTTACCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
      |||  
Db      474 CCTGCTTACCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 533  
  
QY      301 CCAAGATCCCGCAAGCCCTTGGGGGCTCCCAACGGAAGCCATCTTCCGGCGGGATTCT 360  
      |||  
Db      534 CCAAGATCCCGCAAGCCCTTGGGGGCTCCCAACGGAAGCCATCTTCCGGCGGGATTCT 593  
  
QY      361 GATGTGCAACAGTGTGGCGAGCTACGAAACGAGAACGAGCTGTGAGATGCAAT 420  
      |||  
Db      594 GATGTGCAACAGTGTGGCGAGCTACGAAACGAGAACGAGCTGTGAGATGCAAT 653  
  
QY      421 GAGGATGAGAGCACTATCAACAACCAAGCTACCTGTGTGCTTCTGACAGACCCCG 480  
      |||  
Db      654 GAGGATGAGAGCACTATCAACAACCAAGCTACCTGTGTGCTTCTGACAGACCCCG 713  
  
QY      481 GCCACTAGCACTGTGCCCATCAAGCTCTGCACTCAGACACCCCTGGCATCCGACAGT 540  
      |||  
Db      714 GCCACTAGCACTGTGCCCATCAAGCTCTGCACTCAGACACCCCTGGCATCCGACAGT 773  
  
QY      541 GCTTTTCCATGAGTGCATGATGATTAAGTAAAGT 578  
      |||  
Db      774 GCTTTTCCATGAGTGCATGATGATTAAGTAAAGT 811
```

Search completed: November 22, 2004, 12:04:48
Job time : 6720 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 14:06:44 ; Search time 155 Seconds
(without alignments)
469.820 Million cell updates/sec

Title: US-09-597-920b-4_COPY_31_233

Perfect score: 1071

Sequence: 1 RLPGSYDSTSDSLYPRGIQ.....EABEVEERGAPDYENLQELN 203

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1071	100.0	233	2	AA272120 Human LAT
2	683	63.8	242	2	AA272121 Murine LAT
3	132	12.3	710	5	ABB91536 Herbicida
4	128	12.0	374	6	ABP72770 Thiorodox
5	122	11.4	704	7	AD660099 Rat Prote
6	122	11.4	704	7	AD447584 Rat Prote
7	122	11.4	704	7	AD660734 Rat Prote
8	122	11.4	704	7	AD660097 Rat Prote
9	122	11.4	704	7	AD447582 Rat Prote
10	122	11.4	704	7	AD447582 Rat Prote
11	121	11.3	699	7	AD447582 Rat Prote
12	121	11.3	718	8	AAW79156 Human pro
13	121	11.3	742	8	AAW80140 Human pro
14	116	10.8	2058	4	ADP04241 Human col
15	116	10.8	2127	7	ADP09542 Human pol
16	114.5	10.7	838	4	ADP09542 Human nuc
17	114.5	10.7	1000	6	AB256547 Drosophill
18	113.5	10.6	735	6	AB256547 Aspergill
19	113.5	10.6	1186	4	AB256547 Protein e
20	113.5	10.5	1186	6	AB256547 Drosophill
21	112.5	10.5	1186	6	AB256547 Chimeric
22	112	10.5	519	4	ABP71253 Zebrafish
23	112	10.5	519	6	ABP71253 Zebrafish
24	112	10.5	917	6	ADA13333 Amino aci
25	112	10.5	923	6	ABU54617 Human int

ALIGNMENTS

RESULT 1	AA272120	standard; protein; 233 AA.
AC	AA272120;	
DT	14-SEP-1999	(first entry)
DE	Human LAT (linker for activation of T cells) protein.	
KW	LAT; tyrosine kinase; linker for activation of T cell; TCR; human; T-cell receptor; TCR signalling pathway; neoplasia; inflammation; hypersensitivity; allergy; microbial infection; genetic disease; autoimmune disease; graft rejection; modulator.	
OS	Homo sapiens.	
PN	WO9932627-A2.	
PD	01-JUL-1999.	
PF	23-DEC-1998;	98WO-US027400.
PR	23-DEC-1997;	97US-0068690P.
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
XX	Samelson LE, Zhang W;	
XX	WPI: 1999-418926/35.	
XX	N-PSDB; AAX89072.	
XX	Linker for activation of T cell protein used to, e.g. screen for modulators of T cell signaling.	
XX	Claim 1; Fig 7C; 125bp; English.	
XX	The invention relates to a protein tyrosine kinase substrate LAT (linker for activation of T cells) protein. Modulation of interaction between LAT and the T-cell receptor (TCR) affects the TCR signalling pathway. LAT is a substrate for tyrosine kinases and becomes phosphorylated after TCR engagement, resulting in recruitment of other signalling molecules. LAT is used to identify and test (ant)agonists of tyrosine kinase signalling pathways, i.e. modulation of interaction between tyrosine kinase substrates and intracellular ligands or between these ligands and other members of the pathway, including identification of downstream signalling proteins, particularly in immune system cells. These modulators are potentially useful as drugs and diagnostic agents, particularly for	

26	112	10.5	1958	2	AAR60620 Protein f
27	111.5	10.4	216	6	ABP72769 Human nuc
28	111	10.4	1048	2	AAW27277 Human cyt
29	111	10.4	1807	4	AB256597 Recombina
30	111	10.4	2028	4	AB256598 Recombina
31	110.5	10.3	348	4	AB256598 Recombina
32	110	10.3	505	5	AAU77474 EVPR prot
33	110	10.3	571	7	AB081787 Pseudomon
34	109	10.2	966	4	AAW38679 Human pol
35	109	10.2	1013	4	AAW38678 Human pol
36	109	10.2	1025	4	AAW38680 Human pol
37	108.5	10.1	1312	7	ADC31626 Human nov
38	108.5	10.1	774	7	ABP69782 Human nov
39	108.5	10.1	1343	7	AD169279 Human hea
40	108.5	10.1	1714	6	AB225518 Aspergill
41	108.5	10.1	1750	6	AB226118 Aspergill
42	108	10.1	732	4	AB259675 Drosophill
43	108	10.1	2296	4	AB259650 Drosophill
44	107.5	10.0	634	7	ADM04655 Human pro
45	107.5	10.0	1822	5	ABP43899 Nuclear p

CC diseases that involve undesirable cell proliferation, differentiation,
 CC growth or T cell anergy, e.g. neoplasia, inflammation, hypersensitivity/
 CC allergy, microbial infection, metabolic, genetic or autoimmune diseases,
 CC graft rejection. LAT is also used to generate specific antibodies, used
 CC for detection of LAT. Nucleic acid that encodes LAT, or its fragments,
 CC are used to identify homologous sequences in other species; to detect the
 CC LAT gene and as sources of antisense therapeutics. Modulators of LAT are
 CC potentially more specific and less toxic than known immunosuppressants
 CC such as cyclosporin. The present sequence represents the amino acid
 CC sequence of human LAT

XX
 XX Sequence 233 AA;

Query Match 100.0%; Score 1071; DB 2; Length 233;
 Best Local Similarity 100.0%; Pred. No. 6,2e-84;
 Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLPGSYDSTSSSLYPKGIQKRPHTVAPWPPAYPVTSYPLSQPDLPIRSPQPLG 60
 DB 31 RLPGSYDSTSSSLYPKGIQKRPHTVAPWPPAYPVTSYPLSQPDLPIRSPQPLG 90

QY 61 SHRTSSRRDSDGANSVASYENEPACDADDEDYHNPGLVLPDSTPATSTAAPS 120
 DB 91 SHRTSSRRDSDGANSVASYENEPACDADDEDYHNPGLVLPDSTPATSTAAPS 150

QY 121 PALSTPGIRDSAFMESIDYVNVPESGESAASLDGSRVYVNSQLHPGAATPEPAL 180
 DB 151 PALSTPGIRDSAFMESIDYVNVPESGESAASLDGSRVYVNSQLHPGAATPEPAL 210

QY 181 SSQAEEVEEGADPYENLQELN 203
 DB 211 SSQAEEVEEGADPYENLQELN 233

RESULT 2
 AAY27121
 ID AAY27121 standard; protein; 242 AA.

XX
 XX AAY27121;

DT 14-SEP-1999 (first entry)

XX
 XX Murine LAT (linker for activation of T cells) protein.

DE
 XX LAT; tyrosine kinase; linker for activation of T cell; TCR; human;
 KM T-cell receptor; TCR signalling pathway; neoplasia; inflammation;
 KM hypersensitivity; allergy; microbial infection; genetic disease;
 KM autoimmune disease; graft rejection; modulator; mouse.

XX
 XX Mus musculus.

OS
 XX WO9332627-A2.

PN
 XX 01-JUL-1999.

PD
 XX 23-DEC-1998; 98WO-US027400.

PF
 XX 23-DEC-1997; 97US-0068690P.

PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA
 XX Sameleon LE, Zhang W;

PI
 XX WPI; 1999-418926/35.

DR
 XX N-PSDB; AAX89074.

XX
 XX Linker for activation of T cell protein used to, e.g. screen for
 PT modulators of T cell signaling.

XX
 XX Example 2; Fig 7D; 125pp; English.

XX
 XX The invention relates to a protein tyrosine kinase substrate LAT (linker
 CC for activation of T cells) protein. Modulation of interaction between LAT

CC and the T-cell receptor (TCR) affects the TCR signalling pathway. LAT is
 CC a substrate for tyrosine kinases and becomes phosphorylated after TCR
 CC engagement, resulting in recruitment of other signalling molecules. LAT
 CC is used to identify and test (ant)agonists of tyrosine kinase signalling
 CC pathways, i.e. modulation of interaction between tyrosine kinase
 CC substrates and intracellular ligands or between these ligands and other
 CC members of the pathway, including identification of downstream signalling
 CC proteins, particularly in immune system cells. These modulators are
 CC potentially useful as drugs and diagnostic agents, particularly for
 CC diseases that involve undesirable cell proliferation, differentiation,
 CC growth or T cell anergy, e.g. neoplasia, inflammation, hypersensitivity/
 CC allergy, microbial infection, metabolic, genetic or autoimmune diseases,
 CC graft rejection. LAT is also used to generate specific antibodies, used
 CC for detection of LAT. Nucleic acid that encodes LAT, or its fragments,
 CC are used to identify homologous sequences in other species; to detect the
 CC LAT gene and as sources of antisense therapeutics. Modulators of LAT are
 CC potentially more specific and less toxic than known immunosuppressants
 CC such as cyclosporin. The present sequence represents the amino acid
 CC sequence of murine LAT

XX
 XX Sequence 242 AA;

Query Match 63.8%; Score 683; DB 2; Length 242;
 Best Local Similarity 68.9%; Pred. No. 1.5e-50;
 Matches 146; Conservative 16; Mismatches 38; Indels 12; Gaps 6;

QY 2 LPGSYDSTSSSLYPKGIQKRPHTVAPWPPA--YPVTSYPLSQPDLPIRSPQPLG 59
 DB 33 LPVSYDSTSSSLYPKRIILKRPQITVFRTPAVSYPLVTFPPURQDLPIRSPQPLG 92

QY 60 GSHRTSSRRDSDGANSVASYENEPACDADDEDYHNPGLVLPDSTPATSTA 117
 DB 93 GSHRTSSRRDSDGANSVASYENEPACDADDEDYHNPGLVLPDSTPATSTA 151

QY 118 PSAPALSTPGIRDSAFMESIDYVNVPESGESAASLDGSRVYVNSQLHPGAATPE 177
 DB 152 SSAPVPSNPDLGDSAFMESIDYVNVPESGESAASLDGSRVYVNSQLHPGAATPE 210

QY 178 AALSSQEA--EVEEGE-----APPYENLQELN 203
 DB 211 ASVNSQEA--EVEEGE-----APPYENLQELN 242

RESULT 3
 ABB91536
 ID ABB91536 standard; protein; 710 AA.

XX
 XX ABB91536;

AC
 XX 31-MAY-2002 (first entry)

DT
 XX Herbicidally active polypeptide SEQ ID NO 747.

DE
 XX Herbicidal; plant; agriculture; herbicide.

KW
 XX Arabidopsis thaliana.

OS
 XX WO200210210-A2.

PN
 XX 07-FEB-2002.

PD
 XX 28-AUG-2001; 2001WO-EP009892.

PF
 XX 28-AUG-2001; 2001WO-EP009892.

PR
 XX (FARB) BAYER AG.

PA
 XX Tietjen K, Weidler M;

PI
 XX WPI; 2002-269010/31.

DR
 XX Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences

PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms.
 PS Claim 5, SEQ ID NO 747, 261pp + Sequence Listing; English.
 CC The invention relates to identifying target proteins (AB990790-AB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as
 CC herbicides
 CC
 SQ Sequence 710 AA;
 Query Match 12.3%; Score 132; DB 5; Length 710;
 Best Local Similarity 31.6%; Pred. No. 0.015;
 Matches 49; Conservative 6; Mismatches 52; Indels 48; Gaps 6;
 QY 8 STSDSLYPRGIQ-----FKRPHTVAPWPAYPVTSYPLSQDLPPIRSPQPLGGS 61
 DB 40 SPADSSPPALPPLPAVFPSPPTVSSPP--PPLDSSP--PPDLPFPSSPPPPAP 96
 QY 62 -----HRTPSRRSDGANSVASYENEPACEDADEDED----- 95
 DB 97 PPIPIVFPPIIDSPPESTNSPPPEVFPPEPPADEDESPAPPPPEQLPPASSPOGG 156
 QY 96 ----DYHNPGLVLPDSTPATSTAPASAPALSTP 126
 DB 157 PKKPKGHHFG-----PATSPAPSPATSP 182
 RESULT 4
 ABP72770
 ID ABP72770 standard; protein; 374 AA.
 AC ABP72770;
 XX
 DT 23-OCT-2003 (revised)
 DT 11-JUN-2003 (first entry)
 XX
 DE Thioresoxin-hNop140 C-terminal region fusion protein.
 XX
 KW Human; nucleolar phosphoprotein; hNop140; Trx-hNop140C; doxorubicin;
 KW cytosolic; thioresoxin.
 XX
 OS Homo sapiens.
 OS Escherichia coli.
 OS Chimeric.
 XX
 PN WO2003011904-A1.
 PD 13-FEB-2003.
 PF 19-OCT-2001; 2001WO-KR001771.
 PR 30-JUL-2001; 2001KR-00045860.
 PA (KOAD) KOREA ADV INST SCI & TECHNOLOGY.
 XX
 FI Yu Y, Jin Y, Yu J;
 DR WPI; 2003-278388/27.
 DR N-PSDB; AB282080.
 XX
 PT New micro-terminal region protein of human nucleolar phosphoprotein
 PT hNop140, useful for reducing the side effects and improving activity of
 PT an anticancer agent.
 PS Claim 6, Page 31-33; 36pp; English.
 XX

CC The present sequence is the protein sequence of a fusion protein, termed
 CC Trx-hNop140C, comprising thioresoxin linked to the N-terminus of the C-
 CC terminal region (see also ABP72769) of human nucleolar phosphoprotein
 CC hNop140. The C-terminal region of hNop140 was identified by biopanning
 CC as an intracellular target of the anticancer agent, doxorubicin. cDNA
 CC encoding the C-terminal region was amplified by PCR and inserted into
 CC pET28a so as to produce the Trx-hNop140C fusion protein, which was
 CC obtained in large quantities in Escherichia coli transformants. The
 CC fusion protein retains the original form of hNop140 C-terminal region,
 CC and is phosphorylated by casein kinase. The C-terminal region of hNop140
 CC binds to doxorubicin within cells, thus playing an important role in
 CC mediating the cellular toxicity or anticancer activity of doxorubicin.
 CC The hNop140 C-terminal region is expected to be useful for reducing the
 CC side effects of doxorubicin and for improving the anticancer activity of
 CC doxorubicin. (Updated on 23-OCT-2003 to standardise OS field)
 CC
 SQ Sequence 374 AA;
 Query Match 12.0%; Score 128; DB 6; Length 374;
 Best Local Similarity 26.5%; Pred. No. 0.015;
 Matches 53; Conservative 24; Mismatches 87; Indels 36; Gaps 7;
 QY 4 GSYDSTSDSLYPRGIQPKRPHTVAPWPAYPVTSYPLSQDLPPIRSPQPLGSHR 63
 DB 132 GSGDDDDKVLVPRGRSMGNPNS-----SNKPAVTTKSPAYKPAAP----KQVGGGQK 182
 QY 64 TPSSRRSDGANSVASYENEPACEDADEDEDYHNPGLVLPDSTPATSTAPSPAL 123
 DB 183 ILTRRADSSSEESSSEEEKTKK-----WVATTKPATATAKALSLPAK 227
 QY 124 STP-CIRPSAFMSIEDYVNVPEGSGAASLDSREYVNV--OELRPGAKTPPALIS 181
 DB 228 QAPQSSRDS--SSDS-----DSSSESEEEKTKSAVKKKPKQVAGAPSPASAK 277
 QY 182 SQAEVEEGAPDYENLOE 201
 DB 278 KKAESSNSSSSDSSSEEE 297
 RESULT 5
 ADE60099
 ID ADE60099 standard; protein; 704 AA.
 AC ADE60099;
 XX
 DT 29-JAN-2004 (first entry)
 DT
 DE Rat Protein P41777, SEQ ID NO 5998.
 XX
 KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX
 OS Rattus norvegicus.
 OS
 PN WO2003016475-A2.
 PD 27-FEB-2003.
 PF 14-AUG-2002; 2002WO-US025765.
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 PA (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 FI Woolf C, D'Urso D, Befort K, Costigan M;
 DR WPI; 2003-268312/26.
 DR GENBANK; F41777.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page: 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 704 AA;
Query Match 11.4%; Score 122; DB 7; Length 704;
Best Local Similarity 23.2%; Pred. No. 0.11;
Matches 54; Conservative 25; Mismatches 88; Indels 66; Gaps 6;
QY 5 SYDSTSSDLYPRGIGFRRHTVAWMPAYPPVTSYPLSQPDLLP-----RSPQPL 58
Db 372 SSDSDSDSDSDEA-----PAKPVSAATKSLSPAYTPKPAKAAVATPKOP 418
QY 59 GGSHTPSSRRDSDGANSVASYENEPACEDADEDEDYHNPGLVYL.PDSTPATSTAAP 118
Db 419 AGSGQKPKSRKADSSSESSSESSSESSSESSSESSSESSSESSSESSSESSSESS 464
QY 119 SAPALSTP---GIRDSAFSMESIDYVNP-----ESGES 150
Db 465 SLPAKQAPRAGGSSSDSSSSSESSSESSSESSSESSSESSSESSSESSSESSSESS 524
QY 151 AEAALDGSRE-----YVNVSOELHPGAKTPEPALSSQAEVEEAGAPYEN 198
Db 525 SSSSEDSSEEEKKKPKSKATPKPQAGKANGVPASQNGAKGESEEEEDTEQN 577
RESULT 6
ADD47584
ID ADD47584 standard; protein; 704 AA.
XX
AC ADD47584;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein AAA41719, SEQ ID NO 13280.
XX
KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN W02003016475-A2.
XX
PD 27-FEB-2003.

XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-033347P.
XX
PA (GENO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
PI Woolf C, D'urso D, Befort K, Costigan M,
XX
DR WPI; 2003-266312/26.
DR GENBANK; AAA41719.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page: 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 704 AA;
Query Match 11.4%; Score 122; DB 7; Length 704;
Best Local Similarity 23.2%; Pred. No. 0.11;
Matches 54; Conservative 25; Mismatches 88; Indels 66; Gaps 6;
QY 5 SYDSTSSDLYPRGIGFRRHTVAWMPAYPPVTSYPLSQPDLLP-----RSPQPL 58
Db 372 SSDSDSDSDSDEA-----PAKPVSAATKSLSPAYTPKPAKAAVATPKOP 418
QY 59 GGSHTPSSRRDSDGANSVASYENEPACEDADEDEDYHNPGLVYL.PDSTPATSTAAP 118
Db 419 AGSGQKPKSRKADSSSESSSESSSESSSESSSESSSESSSESSSESSSESSSESS 464
QY 119 SAPALSTP---GIRDSAFSMESIDYVNP-----ESGES 150
Db 465 SLPAKQAPRAGGSSSDSSSESSSESSSESSSESSSESSSESSSESSSESSSESSSESS 524
QY 151 AEAALDGSRE-----YVNVSOELHPGAKTPEPALSSQAEVEEAGAPYEN 198
Db 525 SSSSEDSSEEEKKKPKSKATPKPQAGKANGVPASQNGAKGESEEEEDTEQN 577
RESULT 7
ADE60734

ID ADE60734 standard; protein; 704 AA.
XX ADE60734;
AC
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P41777, SEQ ID NO 6646.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KM chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN MO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
XX 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-033347P.
XX
XX (GENO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
DR GENBANK; P41777.
XX
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 704 AA;
Query Match 11.4%; Score 122; DB 7; Length 704;
Best Local Similarity 23.2%; Pred. No. 0.11;
Matches 54; Conservative 25; Mismatches 88; Indels 66; Gaps 6;

QY 5 SYDSTSSDLYPRGIGQKPHVAPWPPAYPVTSYPLSGQDLPIP-----RSPOPL 58
DB 372 SSSSDSSDSSDEDA-----PAKPSATKSPUSKRAVVPKPPAAKAVATPQP 418

QY 59 GGSHTPSSRSDSDGANSVASENEPACEDADEDDYHNPGLVLPDSTPATSTAP 118
DB 419 AGSGGKPGSRKADSSSESSSESEEAATKKS-----VTPPKARVAKAP 464
QY 119 SAPALSTP---GIRDSAFMSGIDDYVAVP-----ESGSS 150
DB 465 SLIPAKQARAGDSSDSESSSESEKKTTPPKPAKKAAGAAYVKKPTPVKKAASSSS 524
QY 151 AEASIDGSR-----YVAVSOELHPGAAKTEBPALSSGPABEEEGADPVEN 198
DB 525 SSSSDSSSESEKPKSKATPKPQAGKANVPASONGKAGKSESEEBDETEGN 577
RESULT 8
ADE60097
ID ADE60097 standard; protein; 704 AA.
XX
XX ADE60097;
AC
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P41777, SEQ ID NO 5996.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KM chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
XX 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-033347P.
XX
XX (GENO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
DR GENBANK; P41777.
XX
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 704 AA:

Query Match 11.4%; Score 122; DB 7; Length 704;
Best Local Similarity 23.2%; Pred. No. 0.11; Mismatches 88; Gaps 6;
Matches 54; Conservative 25; Indels 66; Gaps 6;

5 SYDSTSDSLVPRGIGKRPHTVAPWPAPVPTVPLSQPDLPLP-----RSQPL 58
372 SSSDSDSDSEDEA-----PAKVSATKSPKSPATPKPAPAKAVATKQP 418
QY 59 GGSHTPTSSRRDSDGANSVASYENEPACEDADEDEDYNNPGLVLPDSTPATSTAP 118
DB 419 AGSGQKPKQSRKADSSSSSESSSESESEATKKS-----VTPKAVTAKAP 464
QY 119 SAPPLSTP---GIRDSAFMSIDIDYVNP-----ESGES 150
DB 465 SLPAKQAPRAGGSDSSSESSSESESEKTPPKPPAKKKAAGAVPKTPVKAAAESSSS 524
QY 151 AEASLDGSR-----YNNVSOELHPGAKTEPALSSQAEVEEAGAPYEN 198
DB 525 SSSSESDSSEEEKKPKSKATPKPQAGKANGVPASONGKAGESEEEEDTEQN 577

RESULT 9

ADD47582
ID ADD47582 standard; protein; 704 AA.

XX ADD47582;

DT 29-JAN-2004 (first entry)

XX Rat Protein AAA41719, SEQ ID NO 13278.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX MO2003016475-A2.

XX 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GENO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX Woolf C, D'Urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; AAA41719.

XX Claim 1; Page; 1017p; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence

CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 704 AA:

Query Match 11.4%; Score 122; DB 7; Length 704;
Best Local Similarity 23.2%; Pred. No. 0.11; Mismatches 88; Indels 66; Gaps 6;
Matches 54; Conservative 25; Indels 66; Gaps 6;

5 SYDSTSDSLVPRGIGKRPHTVAPWPAPVPTVPLSQPDLPLP-----RSQPL 58
372 SSSDSDSDSEDEA-----PAKVSATKSPKSPATPKPAPAKAVATKQP 418
QY 59 GGSHTPTSSRRDSDGANSVASYENEPACEDADEDEDYNNPGLVLPDSTPATSTAP 118
DB 419 AGSGQKPKQSRKADSSSSSESSSESESEATKKS-----VTPKAVTAKAP 464
QY 119 SAPPLSTP---GIRDSAFMSIDIDYVNP-----ESGES 150
DB 465 SLPAKQAPRAGGSDSSSESSSESESEKTPPKPPAKKKAAGAVPKTPVKAAAESSSS 524
QY 151 AEASLDGSR-----YNNVSOELHPGAKTEPALSSQAEVEEAGAPYEN 198
DB 525 SSSSESDSSEEEKKPKSKATPKPQAGKANGVPASONGKAGESEEEEDTEQN 577

RESULT 10

ADD47665
ID ADD47665 standard; protein; 704 AA.

XX ADD47665;

DT 29-JAN-2004 (first entry)

XX Rat Protein NP_074060, SEQ ID NO 13361.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX MO2003016475-A2.

XX 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GENO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI: 2003-268312/26.
DR GENBANK; NF_074060.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1, Page: 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNR)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 704 AA;
Query Match 11.4%; Score 122; DB 7; Length 704;
Best Local Similarity 23.2%; Pred. No. 0.11;
Matches 54; Conservative 25; Mismatches 88; Indels 66; Gaps 6;
QY 5 SYDSSSSSLYPRGIGFRRPHTVAPWPAPVPTSYPLSQDPLPIP-----RSQPL 58
DB 372 SSDSDSDSDSEDEA-----PAKPYSAKSPLSKPAVTPKPAKAVATPKQP 418
QY 59 GGSRTSPSSRRDSDGANSVASYENEPACEDADEDDYHNPGLVLPDSTPATSTAP 118
DB 419 AGSGGKPPQSRKADSSSSSEESSESEATKKS-----VTPPKAVYAKKAP 464
QY 119 SAPALSTP---GIRDSAFSMESIDYVNV-----ESGES 150
DB 465 SLPAKQAPRAGDSSSDSSSESEKKTTPKPAKKAAGAAVAPKTPYKKAASESSSS 524
QY 151 AEAALDGSRE-----YNAVSEQLHGAAKTEPALISQEAEEVEEGADYEN 198
DB 525 SSSSESDSESEBKXKPKSKATPKPQAGKANGVAPQNGKAGKESSEEBEDTQCN 577
RESULT 11
AAAM79156
ID AAM79156 standard; protein; 699 AA.
XX
AC AAM79156;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1818.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
OS Homo sapiens.
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
XX
PR 27-APR-2000; 2000US-00560875.
XX
PR 20-JUN-2000; 2000US-00598075.
XX
PR 19-JUL-2000; 2000US-00620325.
XX
PR 01-SEP-2000; 2000US-00654936.
XX
PR 15-SEP-2000; 2000US-00663561.
XX
PR 20-OCT-2000; 2000US-00693325.
XX
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HSE-) HYSEQ INC.
XX
PI Tang YF, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
DR WPI: 2001-476283/51.
DR N-PSDB; AAK52289.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 20; Page 4189-4191; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM79156-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 699 AA;
Query Match 11.3%; Score 121; DB 4; Length 699;
Best Local Similarity 27.3%; Pred. No. 0.11;
Matches 59; Conservative 23; Mismatches 70; Indels 64; Gaps 10;
QY 5 SYDSSSSSLYPRGIGFRRPHTVAPWPPA-----YPPVTSYPLSQDPLPIPSPQ 56
DB 362 SSDSDSDS-----SEDEAPSKPAGTTKNSNNKPAVTTKSPAYKPAAP---KQ 408
QY 57 PLGSHRTSPSSRRDSDGANSVASYENEPACEDADEDDYHNPGLVLPDSTPATSTA 116
DB 409 PVGGGKULTTKADSSSSSESESESEKTK-----MVATTKKATATA 453
QY 117 APSAPALSTP-GIRDSAFSMESIDYVNVPSGSAEALDGSREYVNV--QELHPGAAX 174
DB 454 ALSLPAKQAPQGRDS--SSDS-----DSSSSESESEKTSKSAVKKPKQVAGGAAP 503
QY 175 TEPAL-----SQEAEEVEE-----GAP 194
DB 504 SKPASAKKGAESSSSSDSSSESEBKUKGKGP 539
RESULT 12

AA080140
ID AA080140 standard; protein; 718 AA.
XX
AC AA080140;
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3786.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Dmanac RT, Aundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI: 2001-476283/51.
XX
PT N-PSDB; AAK53273.
XX
PT Nucleic acid encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 20; Page 435; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK5582) and 3666 (AAW80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 718 AA;
XX
Query Match 11.3%; Score 121; DB 4; Length 718;
Best Local Similarity 27.3%; Pred. No. 0.13; Mismatches 70; Indels 64; Gaps 10;
Matches 59; Conservative 23;
XX
QY 5 SYDSTSDSLYPRGIGKRPHTVAWPMPA-----YPVTSYPPISQDPLIPRSPQ 56
DB 381 SSSSDSDS-----SEDDAPSKRAGTTKSSNNKPAVTTKSPAVKPAAP-----KQ 427
QY 57 PLGSHRTPSRRDSDGANSVASYENEPACEDADEDEDYHNGYLVLPDSTPATSTA 116
DB 428 PVGGGQKLLTRKADSSSEESSSESEKTKK-----MVATTKPKATAKA 472

QY 117 APSAPALSTP-GIRDSAFSMESIDYVNVPEGSASASLDSGREYVNVN-QELHPGAK 174
DB 473 ALSLPAKQAPQSGRDS--SSDS-----DSSSESEKTKSKANVKKPKVAGGAP 522
QY 175 TEPAL-----SSQAEVEVEE-----GAP 194
DB 523 SKPASAKKGRAESSSSSDSSSESEKTKKRGSP 558
RESULT 13
ADP04241
ID ADP04241 standard; protein; 742 AA.
XX
AC ADP04241;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human colon specific protein SEQ ID NO:222.
XX
KW human; colon specific nucleic acid; CSNA; colon specific protein; CSP;
KW cytostatic; vaccine; gene therapy; colon cancer.
XX
OS Homo sapiens.
XX
PN WO2004050858-A2.
XX
PD 17-JUN-2004.
XX
PF 04-DEC-2003; 2003WO-US038608.
XX
PR 04-DEC-2002; 2002US-043133P.
XX
XX (DIAD-) DIADEXUS INC.
XX
PI Macina RA, Turner LR, Sun Y, Rodriguez M, Burcham TS;
PI
XX
DR WPI: 2004-480622/45.
XX
XX Novel colon specific protein derived from normal and neoplastic colon
PT cell, useful as vaccine in treating colon cancer and in identifying,
PT diagnosing, monitoring, staging, imaging colon cancer and non-cancerous
PT disease state in colon.
XX
PS Claim 12; SEQ ID NO 222; 655pp; English.
XX
XX The invention relates to a novel colon specific protein (CSP) (I), and
CC the nucleic acid encoding it. A CSP of the invention has cytostatic
CC activity, and may have use in a vaccine, and in gene therapy. The CSP is
CC useful for determining the presence of a colon specific protein in a
CC sample. The nucleic acid encoding the CSP is useful for determining the
CC presence of a colon specific nucleic acid (CSNA) in a sample. The CSP and
CC CSNA are useful for diagnosing or monitoring the presence and metastases
CC of colon cancer in a patient. The method of administering a composition
CC comprising a CSP or CSNA is useful for treating a patient with colon
CC cancer. The CSP and CSNA are useful as vaccine for treating colon cancer
CC and non-cancerous disease states in colon. The present sequence
XX represents a CSP of the invention.
XX
SQ Sequence 742 AA;
XX
Query Match 11.3%; Score 121; DB 8; Length 742;
Best Local Similarity 27.3%; Pred. No. 0.14; Mismatches 70; Indels 64; Gaps 10;
Matches 59; Conservative 23;
XX
QY 5 SYDSTSDSLYPRGIGKRPHTVAWPMPA-----YPVTSYPPISQDPLIPRSPQ 56
DB 405 SSSSDSDS-----SEDDAPSKRAGTTKSSNNKPAVTTKSPAVKPAAP-----KQ 451
QY 57 PLGSHRTPSRRDSDGANSVASYENEPACEDADEDEDYHNGYLVLPDSTPATSTA 116
DB 452 PVGGGQKLLTRKADSSSEESSSESEKTKK-----MVATTKPKATAKA 496
QY 117 APSAPALSTP-GIRDSAFSMESIDYVNVPEGSASASLDSGREYVNVN-QELHPGAK 174

Db 497 ALSLAKAQPOGSRDS--SSDS-----DSSSSESEBEKTSKSAVKKKPKQVAGGAP 546

Qy 175 TEPAL-----SSQBAEVEE-----GAP 194

Db 547 SKPASAKKGAESSNSSSDSDSSEEEEXTLKGKSP 582

RESULT 14

AA97070 ID AAB97070 standard; protein; 2058 AA.

XX AAB97070;

AC 26-JUN-2001 (first entry)

XX 26-JUN-2001 (first entry)

XX Human polypeptide #3 expressed in intraabdominal adipose tissue.

XX Human; intraabdominal adipose tissue; gene expression.

XX Homo sapiens.

XX JP2001008699-A.

XX 16-JAN-2001.

XX 30-JUN-1999; 99JP-00185737.

XX 30-JUN-1999; 99JP-00185737.

XX 30-JUN-1999; 99JP-00185737.

XX (SUMO) SUMITOMO CHEM CO LTD.

XX WPI; 2001-275911/29.

XX Assuming the intraabdominal adipose tissue amount, comprising relating

XX the expression level of a gene expression product to the amount.

XX Claim 1; Page 27-32; 33pp; Japanese.

XX The present sequence is provided in a specification relating to a method

XX for predicting the amount of intraabdominal adipose tissue. The method

XX involves relating the amount of adipose tissue to the expression level of

XX at least one gene transcription product which has a 491 or 2090 residue

XX amino acid sequence and is encoded by a 2376 or 2090 base pair sequence,

XX or which can hybridise to a 2385 or the 2090 base pair sequence under

XX stringent conditions. The method includes a step of deriving the area

XX value of the intraabdominal adipose tissue at the cross section of

XX abdominal navel, from the expression level of a gene transcription

XX product in a sample of the intraabdominal adipose tissue. The method can

XX be used for predicting the amount of intraabdominal adipose tissue

XX irrespective of the amount of subcutaneous adipose tissue. The present

XX sequence is a gene transcription product whose expression level may be

XX measured as part of this method

XX SQ Sequence 2058 AA;

Query Match 10.8%; Score 116; DB 4; Length 2058;

Best Local Similarity 24.9%; Pred. No. 1.3;

Matches 52; Conservative 20; Mismatches 85; Indels 52; Gaps 7;

Qy 5 SYDSSSSSLYPRGQFKRPHTVAPMPAYPVTSYPLSGDILPIRSPQ-PLGSGHR 63

Db 1474 SFGSQQNTSTVP-----PSAPPTTATPLPFTSPFLSGLSSATTPSLPWSAGRS 1476

Qy 64 T-----PSSRRSDGANSVASYENEPACEDADEDDYHNPGLVLPDSTPATST 115

Db 1477 TBEATSSALPEKPGDSVYSASAASLLEBQGSAG-----LPQAPQPSD 1519

Qy 116 AAPSPALSTPGIRDS-----AFSMESIDYVNVPSGSAEASLDGSRRYNVS 165

Db 1520 SVKKEPVLAAQPAVNSGTAASSTSLVALSAEATPATGVPDA--RTEAVPPASSFSV--- 1574

Qy 166 QELHGAAKTEPALSSQBAEVEEBCAP 194

Db 1575 ----PGQTAATTAALSSAGPAVERISRP 1599

RESULT 15

ADF09542 ID ADF09542 standard; protein; 2127 AA.

XX ADF09542;

XX 12-FEB-2004 (first entry)

XX Human nucleoporin 214kDa protein SEQ ID NO:43.

XX human; protein-protein interaction; virucide; cytotoxic; vaccine;

XX human papilloma virus; HPV; cancer.

XX Homo sapiens.

XX WO2003068940-A2.

XX 21-AUG-2003.

XX 14-FEB-2003; 2003WO-US004594.

XX 14-FEB-2002; 2002US-0356911P.

XX (CURA-) CURAGEN CORP.

XX (HOPF) HOFFMANN LA ROCHE INC.

XX Jackson A, Ooi CE, Lewin DA, Cuthill S;

XX WPI; 2003-689668/65.

XX N-PSDB; ADF09547.

XX New purified complex comprising a first polypeptide and a second

XX polypeptide, useful for identifying agents for treating/preventing a

XX condition involving altered level of the complex e.g. human papilloma

XX virus infection, or cancer.

XX Example 3; SEQ ID NO 43; 156pp; English.

XX The invention relates to a novel purified complex comprising a first

XX polypeptide and a second polypeptide, where the polypeptides comprise

XX defined amino acid sequences listed in the specification, and where the

XX first polypeptide binds to the second polypeptide. A complex of the

XX invention has virucide and cytostatic activity, and may have a use as a

XX vaccine. The complex is useful for identifying agents for treating or

XX preventing a conditions involving altered level of the complex, e.g.

XX human papilloma virus (HPV) infection, or cancer. The compositions,

XX antibodies, vectors and methods are useful for treating such diseases.

XX The sequences shown in ADF09500-ADF09583 represent proteins of the

XX invention.

XX SQ Sequence 2127 AA;

Query Match 10.8%; Score 116; DB 7; Length 2127;

Best Local Similarity 24.9%; Pred. No. 1.4;

Matches 52; Conservative 20; Mismatches 85; Indels 52; Gaps 7;

Qy 5 SYDSSSSSLYPRGQFKRPHTVAPMPAYPVTSYPLSGDILPIRSPQ-PLGSGHR 63

Db 1477 SFGSQQNTSTVP-----PSAPPTTATPLPFTSPFLSGLSSATTPSLPWSAGRS 1529

Qy 64 T-----PSSRRSDGANSVASYENEPACEDADEDDYHNPGLVLPDSTPATST 115

Db 1570 TBEATSSALPEKPGDSVYSASAASLLEBQGSAG-----LPQAPQPSD 1572

Qy 116 AAPSPALSTPGIRDS-----AFSMESIDYVNVPSGSAEASLDGSRRYNVS 165

Db 1573 SVKKEPVLAAQPAVNSGTAASSTSLVALSAEATPATGVPDA--RTEAVPPASSFSV--- 1627

Qy 166 QELHGAAKTEPALSSQBAEVEEBCAP 194

Db 1628 ----PCOTAVTAAAISSAGPVAVETSSRP 1652

Search completed: November 21, 2004, 14:16:53
Job time : 168 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 14:11:50 : Search time 39 Seconds
(without alignments)
345,194 Million cell updates/sec

Title: US-09-597-920b-4_COPY_31_233

Perfect score: 1071

Sequence: 1 RLPGSYDSTSSSLVPRGQ.....EAEVEEGAPDYENLQELN 203

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	10.8	2090	US-09-538-092-1081	Sequence 1081, Ap
2	112	10.5	1958	US-07-945-283-2	Sequence 2, Appl
3	111	10.4	455	US-09-270-767-45531	Sequence 45531, A
4	110	10.3	571	US-09-252-991A-30533	Sequence 30533, A
5	106.5	9.9	1298	US-08-690-473-2	Sequence 2, Appl
6	106.5	9.9	1298	US-09-259-821A-2	Sequence 2, Appl
7	106.5	9.9	1298	US-08-843-659-2	Sequence 2, Appl
8	106.5	9.9	1298	US-09-825-288A-2	Sequence 2, Appl
9	105.5	9.9	4019	US-09-854-133-425	Sequence 425, App
10	104.5	9.8	802	US-09-823-240A-2	Sequence 2, Appl
11	104.5	9.8	1219	US-09-344-624-4	Sequence 4, Appl
12	104	9.7	2441	US-08-194-739-2	Sequence 2, Appl
13	104	9.7	2441	US-08-961-739-2	Sequence 2, Appl
14	104	9.7	2441	US-09-514-247A-8	Sequence 8, Appl
15	104	9.7	2441	US-09-686-316-2	Sequence 2, Appl
16	104	9.7	2442	US-09-514-247A-10	Sequence 10, Appl
17	104	9.7	2442	US-09-538-092-1170	Sequence 1370, Ap
18	101.5	9.5	315	US-09-270-767-46043	Sequence 46043, A
19	101	9.4	1048	US-09-171-699-10	Sequence 10, Appl
20	101	9.4	2414	US-08-227-536-2	Sequence 2, Appl
21	101	9.4	2414	US-09-538-092-1189	Sequence 1289, Ap
22	101	9.4	2414	PCT-US95-04682-2	Sequence 2, Appl
23	100.5	9.4	580	US-09-270-767-41648	Sequence 41648, A
24	100.5	9.4	961	US-09-538-092-1231	Sequence 1231, Ap
25	100.5	9.4	1187	US-08-320-559-28	Sequence 28, Appl
26	100.5	9.4	1187	US-08-545-860D-28	Sequence 28, Appl
27	100.5	9.4	1187	PCT-US94-04496-28	Sequence 28, Appl

28	100.5	9.4	1210	1	US-08-320-559-26	Sequence 26, Appl
29	100.5	9.4	1210	3	US-08-545-860D-26	Sequence 26, Appl
30	100.5	9.4	1210	4	US-09-538-092-1179	Sequence 1179, Ap
31	100.5	9.4	1210	5	PCT-US94-04496-26	Sequence 26, Appl
32	98.5	9.2	174	3	US-09-199-637A-63	Sequence 63, Appl
33	98.5	9.2	335	2	US-08-405-175A-6	Sequence 6, Appl
34	98.5	9.2	344	3	US-09-147-236-11	Sequence 11, Appl
35	98.5	9.2	344	4	US-09-522-474-11	Sequence 11, Appl
36	98	9.2	750	3	US-09-165-239A-4	Sequence 4, Appl
37	98	9.2	943	2	US-08-469-537A-107	Sequence 107, App
38	98	9.2	1560	4	US-09-264-512B-2	Sequence 2, Appl
39	97.5	9.1	322	4	US-09-248-796A-17229	Sequence 17229, A
40	97.5	9.1	706	4	US-09-538-092-957	Sequence 957, App
41	97.5	9.1	8921	4	US-08-714-741-32	Sequence 32, Appl
42	97	9.1	282	1	US-07-712-476A-5	Sequence 5, Appl
43	97	9.1	1317	3	US-09-083-521-7	Sequence 7, Appl
44	97	9.1	1970	4	US-09-538-092-1005	Sequence 1005, Ap
45	96.5	9.0	553	3	US-09-083-351-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-538-092-1081
Sequence 1081, Application US/09538092
Patent No. 6753314
GENERAL INFORMATION:
APPLICANT: Glot, Loic
APPLICANT: Mansfield, Traci A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15566-542
CURRENT APPLICATION NUMBER: US/09/538,092
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/127,352
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/178,965
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CuratSeqFormatter Version 0.9
SEQ ID NO 1081
LENGTH: 2090
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION/KEY: (0)...(0)
OTHER INFORMATION: Polypeptide Accession Number P35658
US-09-538-092-1081
Query Match 10.8%; Score 116; DB 4; Length 2090;
Best Local Similarity 24.9%; Pred. No. 0.18;
Matches 52; Conservative 20; Mismatches 85; Indels 52; Gaps 7;
QY 5 SYDSSSLVPRGQFGRHTVAPWPAPVPTVSYPLSQDPLPIRSPQ-PLGSGHR 63
DB 1440 SFGSQNTSTVP-----PSAPPTVATATPLPFTLSGSLSSATPPLPWSAGRS 1492
QY 64 T-----PSRRDSDGANSVASENEPACEDADEDDYHNGYLWVVPDSTPATST 115
DB 1493 TEATSSALPEKPGDSEVSASASILEEQSQAQ-----LPQAPQTSQD 1535
QY 116 AAPSAALSTGIRDS-----AFSMESIDYVNPSSGSAEASLDGSRVNVVS 165
DB 1536 SVKEPVLACPAVNSGTPAASSTSLVALSAEATATTCVPA--RTEAVPPASSFSV--- 1590
QY 166 QELHPGAKTPPALSSQEAEEVEEGAP 194
DB 1591 ----PGQAVTAAAISSAGPVAVETSTTP 1615
RESULT 2
US-07-945-283-2
```

Sequence 2, Application US/07945283
Patent No. 5352596
GENERAL INFORMATION:
APPLICANT: Cheung, Andrew K.
TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
TITLE OF INVENTION: Involving The EPO and LIT Genes
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtie P. Ribando
STREET: 1615 No. 5352596th University Street
CITY: Peoria
STATE: IL
COUNTRY: USA
ZIP: 61604
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ribando, Curtie P
REGISTRATION NUMBER: 27976
TELEPHONE: 309-685-4011 ext.513
TELEFAX: 309-685-4128
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1958 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MODE: TYPE: protein
US-07-945-283-2

Query Match 10.5%; Score 112; DB 1; Length 1958;
Best Local Similarity 35.3%; Pred. No. 0.37;
Matches 30; Conservative 5; Mismatches 26; Indels 24; Gaps 3;

QY 29 PMPAYPVPTVSYPLSQPDLPIPRSPQPLGGS-----HRTSSRRDS 71
Db 483 PSEPPRP-----PLPPPPPPPPPPPPAGSARRRRGGCGGPGGRRRGKRRRA 538
QY 72 DGANSVASYENEPACEDADEDD 96
Db 539 EGTEAAADAEE---EDGDEDEDE 560

RESULT 3
US-09-270-767-45531
Sequence 45531, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 45531
LENGTH: 455
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-45531

Query Match 10.4%; Score 111; DB 4; Length 455;
Best Local Similarity 21.3%; Pred. No. 0.053;
Matches 57; Conservative 23; Mismatches 106; Indels 82; Gaps 9;

QY 4 GSYDSTSSDLYRGIQFKRP-----HTVAPWPPAY---PVT----- 39
Db 165 GSNNNTMMSFFIFLHFNTPTRQQPQQQNVLPANQPTPPFGSAPPAPVASSNNFS 224
QY 40 -----YPLSQPDLPIPRSPQ-----LGSHRTSSRRDS-- 71
Db 225 GQTPMFAAPLNHHPAPWGMGPVVLISIPSPMPASLPWNSPLFKIPPLQAPAKSNDGNQ 284
QY 72 -----DGANSVASYENEPACEDA-----DEDDDYHNPGLVVLPOSTPA 112
Db 285 NDDVDNCPNFSITYSQESQAVANASAMPSCVPHGPADASDDDDDDMD--LVQDDDED 342
QY 113 TSTAAPAPALSTPGIRDSAFMESIDY-----VNPESGSAEASLDGSRBYVNSQ 166
Db 343 TDIPPLIGP---EPEPVKVPKSSDDDLVEPENPTEPEPEPMFEESCAVPTEKSESS 399
QY 167 ELHPGAKTPEPALSSQEAEEVEEGAP 194
Db 400 DHEPSNNSVQAAAPEVNDAAEARTSTP 427

RESULT 4
US-09-252-991A-30533
Sequence 30533, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 30533
LENGTH: 571
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30533

Query Match 10.3%; Score 110; DB 4; Length 571;
Best Local Similarity 22.1%; Pred. No. 0.11;
Matches 54; Conservative 30; Mismatches 92; Indels 68; Gaps 9;

QY 9 TSSDSLVRGIGFKRPHTVAPWPPAY-----PVTSPPLSQPDLPIPRSPQ 56
Db 315 TPTPTVSPSG-----SVARQAPAVSARVASTOAREPASVSPVDEBPLVSSHPQ 367
QY 57 PLGSHRTPS-----SRSDGANSV-ASVENEPACEDADE--DDYHNPGLVVL 108
Db 368 IAGRTHERPQGGGFPKTAIEVAISTQASVQSPPTGAGRGGERQPG--ETDPS 425
QY 109 STPATSTAPASAPALSTPGIR-----DSAFMESIDYVNVNPESESA 151
Db 426 ALPEDDAPVPLPAMQPGRLVARRLASSGSRPLPLADLARRLDVAGRIQVAAASH 485
QY 152 EASLDGSRBYVNSQELHPGAKTPEPALSSQEAEEVEEGAD-----YENT 199
Db 486 AARL-----QVRLPOLGAVEVOVLHGQQLQVEISASPSGLAFLQAGARELLERL 535
QY 200 QELN 203
Db 536 QRLH 539

RESULT 5
US-08-690-473-2
Sequence 2, Application US/08690473
Patent No. 5876923
GENERAL INFORMATION:

```

; APPLICANT: Leopardi, Rosario
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 AS AN
; TITLE OF INVENTION: INHIBITOR OF APOPTOSIS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,473
; FILING DATE: 26-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: ARCD:239
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1298 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-690-473-2

Query Match          9.9% Score 106.5; DB 2; Length 1298;
Best Local Similarity 24.0%; Pred. No. 0.66;
Matches 49; Conservative 19; Mismatches 73; Indels 63; Gaps 7;

QY 52 PRSPQPLGSHRTSPSSRRSDSGANSVASYENEE-----PAC 87
DB 9 PGSPGPTDGPPTPSPDRDERGALCMGA-ETEGGDDPDHPDHPDLDARDDRARA 67
QY 88 -EDADEDDYHNPGLVVL---PDSTPATSTAPSPALSTPGIRDSAFMESIDYV 142
DB 68 GTDAGEDDAGDAVSPQLALLASWVEAVRTTPTPDPAASPPRTPAFRADDDGDEYDPA 127
QY 143 N-----VESGESASASIDGSRVYVNGQLHP-----G 171
DB 128 DAAGDRAPARGRERAPLRGA--YDPPTDRLSPPPPAOPRRRRRGRWRPSASTSSDSG 185
QY 172 AAKTEPALSSQAEVEVEEGAPD 195
DB 186 SSSSSASASSSSSDDEDDEDDGND 209

RESULT 6
US-09-259-821A-2
; Sequence 2, Application US/09259821A
; Patent No. 6210926
; GENERAL INFORMATION:
; APPLICANT: LEOPARDI, ROSARIO
; APPLICANT: ROIZMAN, BERNARD
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
; FILE REFERENCE: ARCD:317
; CURRENT APPLICATION NUMBER: US/09/259,821A
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/690,473
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
```

```

; LENGTH: 1298
; TYPE: PRT
; ORGANISM: HERPES VIRUS, TYPE 1
;
US-09-259-821A-2

Query Match          9.9% Score 106.5; DB 3; Length 1298;
Best Local Similarity 24.0%; Pred. No. 0.66;
Matches 49; Conservative 19; Mismatches 73; Indels 63; Gaps 7;

QY 52 PRSPQPLGSHRTSPSSRRSDSGANSVASYENEE-----PAC 87
DB 9 PGSPGPTDGPPTPSPDRDERGALCMGA-ETEGGDDPDHPDHPDLDARDDRARA 67
QY 88 -EDADEDDYHNPGLVVL---PDSTPATSTAPSPALSTPGIRDSAFMESIDYV 142
DB 68 GTDAGEDDAGDAVSPQLALLASWVEAVRTTPTPDPAASPPRTPAFRADDDGDEYDPA 127
QY 143 N-----VESGESASASIDGSRVYVNGQLHP-----G 171
DB 128 DAAGDRAPARGRERAPLRGA--YDPPTDRLSPPPPAOPRRRRRGRWRPSASTSSDSG 185
QY 172 AAKTEPALSSQAEVEVEEGAPD 195
DB 186 SSSSSASASSSSSDDEDDEDDGND 209

RESULT 7
US-08-843-659-2
; Sequence 2, Application US/08843659
; Patent No. 6218103
; GENERAL INFORMATION:
; APPLICANT: Leopardi, Rosario
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS US3 AND ICP4 AS
; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/843,659
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: ARSB:519
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1298 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-843-659-2

Query Match          9.9% Score 106.5; DB 3; Length 1298;
Best Local Similarity 24.0%; Pred. No. 0.66;
Matches 49; Conservative 19; Mismatches 73; Indels 63; Gaps 7;

QY 52 PRSPQPLGSHRTSPSSRRSDSGANSVASYENEE-----PAC 87
```

Db 9 PGSPGTPPPPTSPDRDERGALWGGA-ETEEGGDDPDHDPHDLDDARRDGRAPAA 67
Qy 88 -EDADEDEDYHNGYLVV- -PDSTPATSTAPASAPALSTGIRDSAFMESIDYV 142
Db 68 GTDAGEDAGDAVSPROLALLASWVEAVRTIPTDPAPASPPRTAPFADDDGDEYDAA 127
Qy 143 N-----VPEGESAEASLDGSEYVNVSOELHP-----G 171
Db 128 DAAGDRAPARGREAREAPLRGA--YDPDTRLSPPRPAOPRRRRHGRWRPESASTSSDSG 185
Qy 172 AAKTEPALSSQAEVEVEEGAPD 195
Db 186 SSSSSSSASSSSSSDEDEDDGND 209

RESULT 8
US-09-825-288A-2
Sequence 2, Application US/09825288A
Patent No. 6723511
GENERAL INFORMATION:
APPLICANT: LEOPARDI, ROSARIO
APPLICANT: ROIZMAN, BERNARD
TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
FILE REFERENCE: ABCD:3170SC1
CURRENT APPLICATION NUMBER: US/09/825,288A
CURRENT FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: 09/2259,821
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: 08/690,473
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1298
TYPE: PRT
ORGANISM: HERPES VIRUS, TYPE 1
US-09-825-288A-2

Query Match 9.9%; Score 106.5; DB 4; Length 1298;
Best Local Similarity 24.0%; Pred. No. 0.66; Indels 63; Gaps 7;
Matches 49; Conservative 19; Mismatches 73;
Qy 52 PRSPQPLGSHRTSPSRSDGANSVASYENEE-----PAC 87
Db 9 PGSGPTDGPPTSPDRDERGALWGGA-ETEEGGDDPDHDPHDLDDARRDGRAPAA 67
Qy 88 -EDADEDEDYHNGYLVV- -PDSTPATSTAPASAPALSTGIRDSAFMESIDYV 142
Db 68 GTDAGEDAGDAVSPROLALLASWVEAVRTIPTDPAPASPPRTAPFADDDGDEYDAA 127
Qy 143 N-----VPEGESAEASLDGSEYVNVSOELHP-----G 171
Db 128 DAAGDRAPARGREAREAPLRGA--YDPDTRLSPPRPAOPRRRRHGRWRPESASTSSDSG 185
Qy 172 AAKTEPALSSQAEVEVEEGAPD 195
Db 186 SSSSSSSASSSSSSDEDEDDGND 209

RESULT 9
US-09-854-133-425
Sequence 425, Application US/09854133
Patent No. 6759508
GENERAL INFORMATION:
APPLICANT: Lodes, Michael J.
APPLICANT: Mohamach, Radoch
APPLICANT: Henderson, Robert A.
APPLICANT: Benson, Darin R.
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
FILE REFERENCE: 210121.475C10
CURRENT APPLICATION NUMBER: US/09/854,133

CURRENT FILING DATE: 2001-05-11
NUMBER OF SEQ ID NOS: 735
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 425
LENGTH: 4019
TYPE: PRT
ORGANISM: Homo sapiens
US-09-854-133-425

Query Match 9.9%; Score 105.5; DB 4; Length 4019;
Best Local Similarity 27.0%; Pred. No. 3.8; Indels 49; Gaps 11;
Matches 61; Conservative 19; Mismatches 97;
Qy 3 GSYDSTSSDGLYPRGIGFRRPHT-----VAPW-----PRAPVPTSTYF-----PLSQP 46
Db 1218 PGTRPRTTVDP-YSQDQGTREPTQTDILFTVTNQHNSDPVHAPGTRPGISVPSQ 1276
Qy 47 DLLPFRSPQPLGSHRT-PSRRSDGANSVASYENEEPCADADEDEDYHNGYLVV 105
Db 1277 PATPRPISGFRSSWTRVLPNODPFLQAA--QNRGPA-----PGPLVR 1322
Qy 106 LPDSTPATSTAPASAPALSTGIRDSAFMESIDYVNVPEGESAEASLDGSEYVNV 165
Db 1323 PPDT--CSQTPRPPGELSDTFSRVSPPAAR--DPYDQSPMTRSSQSDFTGISTAHVA 1378
Qy 166 QELHPG-----AAKTEPALSSQAEVEVE-----EGAPDYEN 198
Db 1379 DQRRPGESEGFCASSNSPMHISQDQFSGVSGQLGCPVTSVDTTON 1424

RESULT 10
US-09-823-240A-2
Sequence 2, Application US/09823240A
Patent No. 6716597
GENERAL INFORMATION:
APPLICANT: Frank B. Gertler
APPLICANT: James E. Bear
APPLICANT: Jurgen Loureld
APPLICANT: Joseph Loureld
TITLE OF INVENTION: Methods and Products for Regulating Cell
FILE REFERENCE: M00656.70064.US
CURRENT APPLICATION NUMBER: US/09/823,240A
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 60/194,564
PRIOR FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 802
TYPE: PRT
ORGANISM: Mus musculus
US-09-823-240A-2

Query Match 9.8%; Score 104.5; DB 4; Length 802;
Best Local Similarity 24.0%; Pred. No. 0.52; Indels 77; Gaps 9;
Matches 60; Conservative 17; Mismatches 96;
Qy 4 GSYDSTSSDGLYPRGIGFRRPHTVAPWPPAYPPVTSYPLISQDPLFT-----PRSP 55
Db 422 GSIDSVT-----YVSPPTSPGPAAPPPPPPPPPPPPLPLPLPLSLHSGSQASP 477
Qy 56 QPLGSHRTSPSRD---SPGANSVASYEN-----EPPACEDDEDDED-----96
Db 478 PGTPLASTPSKPSVLPSPSAGAPASAEPLPLPELIDSSASBPGLQAAQPAESPPOG 537
Qy 97 -----YHNGYLVV-----PDSTPATSTA-----116
Db 538 LVIGPPAPPPPPPLPSGAPASALPPPPGPPPPPLPSTGPPPPPPPPPPPPPPPPPP 597
Qy 117 APSAPALSTGIRDSAFMESIDYVNVPEGESAEASLDGSEYVNVSOELHPGAATE 176
Db 598 PPAPPLPAPGI-----FSGSTSD--NRPLTGLAAIAGAKLHKVSRVDESGSPGCGNTG 651

QY 177 PAALSSOEA 186
 DB 652 SVSLASSKAD 661

RESULT 11

US-09-344-624-4
 ; Sequence 4, Application US/09344624
 ; Patent No. 6753154
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Hui-Mei
 ; APPLICANT: Bissell, Mina
 ; TITLE OF INVENTION: HUMAN AZ-1 GENE, VARIANTS THEREOF AND EXPRESSED GENE
 ; TITLE OF INVENTION: PRODUCTS
 ; FILE REFERENCE: 2960.44 (HV)
 ; CURRENT APPLICATION NUMBER: US/09/344,624
 ; CURRENT FILING DATE: 1999-06-25
 ; EARLIER APPLICATION NUMBER: 60/090,747
 ; EARLIER FILING DATE: 1998-06-26
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 1219
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-344-624-4

Query Match 9.8%; Score 104.5; DB 4; Length 1219;
 Best Local Similarity 22.5%; Pred. No. 0.92;
 Matches 64; Conservative 28; Mismatches 73; Indels 119; Gaps 16;

QY 5 SYDS-----TSSDSLYPRGICQ---FKRPHYAP--WPPAYP-----VTSYPL 43
 DB 280 SYHSDVQGVSTDLIAQSSDSEBAFETPESSTPVKAPAPAPPPPEVYPEPEVSTOPPP 339
 QY 44 SOP-----DLPLI---FRSPQPLGSHRTPPS-----RRSDGANSYA 78
 DB 340 EEPGCGSTVPVPPGKSDSVGSPFRPPSHFSAVPBEDQPIASSGYNLDFDNIELVD 399
 QY 79 SYNEBPACEDADEDE-----DDYH-----NPGYLV 104
 DB 400 TFOGLEPASPADKONGKVTNRKSTDSVPIKSTLSLSQASDFGASSGNGPEAVA 459
 QY 105 VLFPD-----STPATS-----TAAPAPAL-----STGIRDASMSIDYVNP 145
 DB 460 LAPPAVSTGSSSASTLKRTKPRPPSLKKQTKKPTETPVKE---TOGEPDESLVP 516
 QY 146 ESGEASASLDGSRVNVVSOELHPGAKTE---PAALSSOEA 186
 DB 517 SGE-----NLASRTKTESAKTBGSPALLETPE 546

RESULT 12

US-08-194-468-2
 ; Sequence 2, Application US/08194468
 ; Patent No. 5750336
 ; GENERAL INFORMATION:
 ; APPLICANT: Montminy, Marc R.
 ; TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF
 ; TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN
 ; TITLE OF INVENTION: RESPONSIVE GENES
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
 ; STREET: 444 South Flower Street, Suite 2000
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 90071
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/194,468
 ; FILING DATE: 10-FEB-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Reiter, Stephen E.
 ; REGISTRATION NUMBER: 31,192
 ; REFERENCE/DOCKET NUMBER: P41 9672
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619)-546-4737
 ; TELEFAX: (619)-546-9392
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2441 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-194-468-2

Query Match 9.7%; Score 104; DB 1; Length 2441;
 Best Local Similarity 25.8%; Pred. No. 2.6;
 Matches 54; Conservative 18; Mismatches 79; Indels 58; Gaps 10;

QY 3 PGSDSTSSDLYPRGICQFKRPHYAPWP--PAYPVTSYPLGQDILLPIRSPQPIGG 60
 DB 904 PGSVPSAAQTOSTPT-VQAAQAQVTPQPTVPQPSVATQSSQQ--PPVHTQPPG- 959
 QY 61 SHRPSSRRSDGANSVSYNEBPACEDADEDEDYNNPYLVLPSTPATSTAAPSA 120
 DB 960 ---TPLSQ-----AAASIDNRVP-----TSTYTSATS 985
 QY 121 PALSTPGIRDAFMSIDYVNPESGES-----AEASLDGSRVNVVSOELHPGA 172
 DB 986 SQQCPQDVMLEMTETVQTDAB-PEPTESKGEPRSEKMEEDLQSSQ---VKETDTTE 1041
 QY 173 AKTEPALSSQ-----EAEEVEEGAPD 195
 DB 1042 QKSEPEVEERKPEVKEAKEEBENSSND 1070

RESULT 13

US-08-961-739-2
 ; Sequence 2, Application US/08961739A
 ; Patent No. 6063583
 ; GENERAL INFORMATION:
 ; APPLICANT: Montminy, Marc R.
 ; TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
 ; FILE REFERENCE: SALK1650-1
 ; CURRENT APPLICATION NUMBER: US/08/961,739A
 ; CURRENT FILING DATE: 1997-10-31
 ; EARLIER APPLICATION NUMBER: US 194,468
 ; EARLIER FILING DATE: 1994-02-10
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 2441
 ; TYPE: PRT
 ; ORGANISM: Mus
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)...(2441)
 ; OTHER INFORMATION: Xaa = Any Amino Acid
 ; US-08-961-739-2

Query Match 9.7%; Score 104; DB 3; Length 2441;
 Best Local Similarity 25.8%; Pred. No. 2.6;
 Matches 54; Conservative 18; Mismatches 79; Indels 58; Gaps 10;

QY 3 PGSDSTSSDLYPRGICQFKRPHYAPWP--PAYPVTSYPLGQDILLPIRSPQPIGG 60
 DB 904 PGSVPSAAQTOSTPT-VQAAQAQVTPQPTVPQPSVATQSSQQ--PPVHTQPPG- 959

```

QY 61 SHRTSSNRSDSGANSVANSYENEEBACADADEDEDYHNPGLVYLPSTPATSTAAASA 120
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 960 ---TFLSQ-----AAASIDNRVD-----TSTVTSATSS 985
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 121 PALSTPGIRDSAFSMESIDYVNVSEGES-----AASLDGSRXYVNVSCGLHGA 172
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 986 SQQSPEDVPMLEAKTEVOTDDAE-DEPLEBSKCEPRSEMMEDLQSSQ---VKETDTTE 1041
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 173 AKTEPALASQ-----EAEVEEBGARD 195
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1042 QKSEPMVEYBEKKPEYVKAKEEENSSND 1070
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

```

RESULT 14
US-09-514-247A-8
: Sequence 8, Application US/09514247A
: Patent No. 6365361
: GENERAL INFORMATION:
: APPLICANT: TANABE SEIYAKU CO. LTD.
: APPLICANT: TANIGUCHI, Tomoyasu
: APPLICANT: MIZUKAMI, Junko
: TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PPR
: FILE REFERENCE: TANIGUCHI-6
: CURRENT APPLICATION NUMBER: US/09/514,247A
: CURRENT FILING DATE: 2000-02-28
: PRIOR APPLICATION NUMBER: PCT/JP98/03734
: PRIOR FILING DATE: 1998-08-24
: PRIOR APPLICATION NUMBER: JP231084/1997
: PRIOR FILING DATE: 1997-08-27
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO: 8
: LENGTH: 2441
: TYPE: PRT
: ORGANISM: mouse
US-09-514-247A-8

```

	Query Match	Similarity	9.7%	Score 104	DB 3	Length 2441
Match	Local	54.1	25.8%	Pred. NO. 2.6		
	Conservative	54	18	Mismatches	79	Indels 58; Gaps 10.
Qy	3	PGSYDSTSSDLSLYPEGRIQFKRHVAAMP--PAYPVVTSYELPSQPDLLPIPRSDQPIGG	60			
Db	904	PGSVPSAAQOTGSTPR-VQAAQAQVTPQGQGVQPSVATPQSSQGC--PFRVHTQPG-	959			
Qy	61	SHRTSSRRDDGANSVASYENEPACEDADEDDYNNPGVLVLPSTPATSTAAASA	120			
Db	960	--TFLSQ-----AAASIDNRVP-----TSTVTSASTS	985			
Qy	121	PALSTPGIRDSAFSMESIDDVNVNPESGES-----AASLDGSRVNVNVSQELHFGA	172			
Db	986	SQQSPDPVPMLEMKTEVQTDAE-PEPESKKEPRSEMMEDLQSSGC--VKERTDTTE	1041			
Qy	173	AKTEPALSSQ-----EAEEVEEGCAPD	195			
Db	1042	QKSEMEVEEKKPEYKVAKEEENSSND	1070			

RESULT 15
US-09-686-316-2
Sequence 2, Application us/09686316
Patent No. 6646115
GENERAL INFORMATION:
APPLICANT: MONTMAY, Marc R.
TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
FILE REFERENCE: SAL1650-1
CURRENT APPLICATION NUMBER: US/09/686,316
CURRENT FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: US/08/961,739
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: US 194,468
PRIOR FILING DATE: 1994-02-10

```

? NUMBER OF SEQ ID NOS: 4
? SOFTWARE: FastSeq for Windows Version 4.0.0
? SEQ ID NO 2
? LENGTH: 2441
? TYPE: PRT
? ORGANISM: Mus
?
? FEATURE:
? NAME/KEY: VARIANT
? LOCATION: (1)...(2441)
? OTHER INFORMATION: Xaa = Any Amino Acid
US-09-686-316-2

```

Query match

9.7%; Score 104; DB 4; Length 2441;

Best Local Similarity 25.8%; Pred.NO. 2.6;
Matches 54; Conservative 18; Mismatches 79; Indels 58; Gaps 10;

[illegible]

Search completed: November 21, 2004, 14:21:37
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 14:20:19 ; Search time 141 Seconds
(without alignments)
509.843 Million cell updates/sec

Title: US-09-597-920B-4_COPY_31_233

Perfect score: 1071

Sequence: 1 RLPSYDSTSSDLSLPRGIQ.....EAEVEBEGAPDYENLQELN 203

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	123.5	11.5	311	US-10-437-963-174234	Sequence 174234,
2	120.5	11.3	468	US-10-437-963-146700	Sequence 146700,
3	114.5	10.7	1000	US-10-128-714-3305	Sequence 3305, Ap
4	114	10.6	762	US-10-437-963-131253	Sequence 131253,
5	113.5	10.6	735	US-10-282-122A-69392	Sequence 69392, A
6	112.5	10.5	1177	US-10-193-692-4	Sequence 4, Appli
7	112.5	10.5	1186	US-10-193-692-2	Sequence 2, Appli
8	112	10.5	430	US-10-739-930-9638	Sequence 9638, Ap
9	112	10.5	519	US-10-113-794A-2	Sequence 2, Appli
10	112	10.5	519	US-10-428-487-14	Sequence 14, Appli
11	112	10.5	519	US-10-258-371B-28	Sequence 28, Appli
12	112	10.5	923	US-10-114-270-152	Sequence 152, App
13	111	10.4	169	US-10-425-115-224853	Sequence 224853,

14	110	10.3	346	US-10-437-963-184670	Sequence 184670,
15	110	10.3	584	US-10-156-761-12405	Sequence 12405, A
16	109.5	10.2	340	US-10-437-963-181586	Sequence 181586,
17	109.5	10.2	1151	US-09-825-751A-79	Sequence 79, Appli
18	109.5	10.2	1240	US-10-369-493-4031	Sequence 4031, Ap
19	109	10.2	1678	US-10-437-963-138217	Sequence 138217,
20	108.5	10.1	1343	US-10-408-765A-1085	Sequence 1085, Ap
21	108.5	10.1	1714	US-10-128-714-3176	Sequence 3176, Ap
22	108.5	10.1	1750	US-10-128-714-8176	Sequence 8176, Ap
23	108	10.1	509	US-10-437-963-106493	Sequence 106493,
24	107.5	10.0	634	US-10-108-260A-3340	Sequence 3340, Ap
25	106.5	9.9	102	US-09-739-907-58	Sequence 58, Appli
26	106.5	9.9	102	US-09-938-671-58	Sequence 58, Appli
27	106.5	9.9	180	US-09-739-907-106	Sequence 106, App
28	106.5	9.9	180	US-09-938-671-106	Sequence 106, App
29	106.5	9.9	420	US-10-425-115-184837	Sequence 184837,
30	106.5	9.9	1298	US-09-825-288A-2	Sequence 2, Appli
31	106	9.9	1372	US-10-437-963-165419	Sequence 165419,
32	106	9.9	1790	US-10-437-963-138166	Sequence 138166,
33	105.5	9.9	283	US-09-864-761-34492	Sequence 34492, A
34	105.5	9.9	553	US-10-437-963-147077	Sequence 147077,
35	105.5	9.9	4019	US-09-738-973-425	Sequence 425, App
36	105.5	9.9	4019	US-09-854-133-425	Sequence 425, App
37	105.5	9.9	4019	US-10-144-649A-425	Sequence 425, App
38	105	9.8	1164	US-10-437-963-165343	Sequence 165343,
39	105	9.8	1345	US-10-437-963-165375	Sequence 165375,
40	104.5	9.8	497	US-09-764-864-1314	Sequence 1314, Ap
41	104.5	9.8	534	US-09-764-864-861	Sequence 861, App
42	104.5	9.8	678	US-10-437-963-165341	Sequence 165341,
43	104.5	9.8	802	US-09-823-240-2	Sequence 2, Appli
44	104.5	9.8	1059	US-10-425-115-222100	Sequence 222100,
45	104	9.7	553	US-10-771-931-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1
US-10-437-963-174234
; Sequence 174234, Application US/10437963
; Publication No. US20040123343A1
GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 174234
LENGTH: 311
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(311)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_72195C.1.pep
US-10-437-963-174234
Query Match 11.5%; Score 123.5; DB 16; Length 311;
Best Local Similarity 23.1%; Pred. No. 0.15;
Matches 46; Conservative 30; Mismatches 84; Indels 39; Gaps 6;
QY 30 WPAVPTSTYPLSPD---LPLRPSPQLGSHRPPSSRRSDGANSVASYENEDPA 86

```

Db      53 WAPAPLTRSWADVEDDDDDXYFATTAPPRPVWGTHHHAADHDDHDEQALQLESE 112
Qy      87 CEDADEDEDYHNPGYLVLPDSTPA-----TSTAAPSALSTPG----- 127
Db      113 DEVEDDDDEDHEH---ETEDATPAEPANMKAAAPAPKPTERQSKELKKKELEEL 168
Qy      128 ---TRDSAFMESIDYVNVPESGESASLDGSRVYVNSQELHPGAKTPEPALSSQE 184
Db      169 DAILAELELSKSNNDQANFNNGKGAQAADGE-----NKEGAPAPAESK-----SSKK 218
Qy      185 AEEVEEGAPDYENLOELN 203
Db      219 KKAKDKSAKEAKETQELN 237

RESULT 2
US-10-437-963-146700
; Sequence 146700, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 146700
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(468)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_4729C.1.pep
US-10-437-963-146700

Query Match      11.3%; Score 120.5; DB 16; Length 468;
Best Local Similarity 24.9%; Pred. No. 0.42;
Matches 54; Conservative 28; Mismatches 82; Indels 53; Gaps 7;

Qy      1 RLPGSYSTSDSLYPRGIQFKRHTVAPWPAVPVTSYPLSQDULPIRSPQPLG 60
Db      6 RRPDSG-----LARGQWARRREVARMLEBLAPVTGTSETTXEMPKRVK---- 53
Qy      61 SHRTSPSRDSDGANSVASYENEPACDADDEDYHNPGYLVLPDSTPATSTAAPSA 120
Db      54 --EKVVRKESDADGPMVAEEGAPSAVAADGE-----GQASQSPSAPAPQPS-SA 104
Qy      121 PALSTPGIRDSAFMESIDYVNVPESGESABA-----SLDGSREYVNSQELHPGAKT 175
Db      105 PATS-----VQVPNTADVAKAAVAVALCTRAENLSTNOLVVOAAPS 147
Qy      176 EPAA-----LSQEAEEVEEGAPDYENLOEL 202
Db      148 QPAAPTALAVVQAQISLDPAQAQAEADWEARRQNMTRL 184

RESULT 3
US-10-128-714-3305
; Sequence 3305, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128, 714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3305
; LENGTH: 1000
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3305

Query Match      10.7%; Score 114.5; DB 14; Length 1000;
Best Local Similarity 23.3%; Pred. No. 3.1;
Matches 49; Conservative 32; Mismatches 98; Indels 31; Gaps 8;

Qy      7 DSTSDSLYPRGIQFKRHTVAPWPAVPVTSYPLSQDULPIRSPQPLGSHRTPS 66
Db      293 ESAPSESVAAEPASABEP-----PPADEPAEPPEPAESGAPAVEVEEP-APQDETIV 345
Qy      67 SRRDSDGANSVASYENEPACDADDEDYHNPGYLVLPDSTPA-TSTAAPSAPALS 124
Db      346 QKAPAAEQSTADYVTEPEPAVVEASAKP-----VAEPTPDGKSPVEBAVVE 396
Qy      125 TPGIRDSAFMES-IDYVNVPESGESASLDGSRVYVNSQOE-----LHPGAK 174
Db      397 APPAESASGSESTPYEEAAPTEESAESAPEAEATESISVGEPAADAEBGLRESAH 456
Qy      175 TEPAALSSQEAEE---VEEGAPDYENLOE 201
Db      457 EEPAPVEDADPAEPAVEEPPAAEESPSVED 486

RESULT 4
US-10-437-963-131253
; Sequence 131253, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 131253
; LENGTH: 762
; TYPE: PRT
; ORGANISM: Oryza sativa

```



```
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_3336C.1.pap
US-10-437-963-131253

Query Match
Best Local Similarity 10.6%; Score 114; DB 16; Length 762;
Matches 52; Conservative 21; Mismatches 65; Indels 54; Gaps 10;

QY 29 PMPAPVPTYSYR-----LSQPDLPPIPSRPPGLSGHRTSSRRSDGANSVASYENE 84
DB 42 PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP
QY 85 PACEDADEDEDY-----HNPGLVLPDSTPATSTA-----APSAPALSTPGIR 129
DB 90 EAEIEEEDDEDHLDRRRLRLHRRPQVPLVSSPAGCTVTPQQPPPPPPPPPPPPPPPP 149
QY 130 --DSAPSE---SI---DDYVNPESGESASASLDGSEYVNVSGELHGAATKTEPA 179
DB 150 TWDFPFMSDEGMASIAPDDEIIQEBE-----DEKTVPASPPRPPPLSPPTPVAA 198
QY 180 LSSQAEVEEER 191
DB 199 PAPQ--EEFEER 208

RESULT 5
US-10-282-122A-69392
Sequence 69392, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangshu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haeselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Twawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELTRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 69392
LENGTH: 735
TYPE: PRT
ORGANISM: Pseudomonas syringae
```

```
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (445T..(445)
OTHER INFORMATION: X-any amino acid
US-10-282-122A-69392

Query Match
Best Local Similarity 10.6%; Score 113.5; DB 15; Length 735;
Matches 55; Conservative 22; Mismatches 80; Indels 49; Gaps 11;

QY 23 RPHVAPMPAPVPTYSYRPPPLSQPDLPPIPSRPP-----LGSHRTSSRRSDG----- 73
DB 413 QPPVAPAPAPVAPQPAKAP--APQIKPEP--EPQXTQACGRRRSAAVERVESAAGRKA 469
QY 74 -----ANSVASYENE-----EPACEDADEDEDYHNPGLVLPDSTPATSTA 116
DB 470 CRARARARARASCAVEAEPOPEVAPPEVLETVEQPD-----LTPMAPAPAPSP 520
QY 117 APSAP--ALSTPGIRSAFS---MESIDD--YVNPESGESASASLDGSEYVNVSGELHP 170
DB 521 VPDAPQAPSPPEVEEQVTPAMLEAI PDSAYUSAPMDXDEPPAD--DIVEPDIDIF 577
QY 171 GAAK--TEPAAALSSQAEVEEER 194
DB 578 ASYSYLDLAHESVVELEAVEPEPAP 603

RESULT 6
US-10-193-692-4
Sequence 4, Application US/10193692
Publication No. US20030074682A1
GENERAL INFORMATION:
APPLICANT: Exelixis Deutschland GmbH
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF A NOVEL TELBOST POTASSIUM
FILE REFERENCE: ARO2-005C
CURRENT APPLICATION NUMBER: US/10/193,692
CURRENT FILING DATE: 2002-07-11
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 1177
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION:
US-10-193-692-4

Query Match
Best Local Similarity 10.5%; Score 112.5; DB 14; Length 1177;
Matches 49; Conservative 20; Mismatches 61; Indels 25; Gaps 10;

QY 27 VAPMPAPVPTYSYRPPPLSQPDLPPIPSRPP--GSHRTSSRRSDGANSVASYENE 84
DB 1039 MAPVPAPVAV--SPDPLAHF--VP--PAHPTSLYTAHNTTPSLQITD-----ASSPKS 1089
QY 85 PACEDADEDEDYHNPGLVLPDST-----PATSTAPSA--PALSTGIRDSATSMESI 138
DB 1090 PDVDSLKERKSPDLSGSHLTVASTDTMSKSPETELSVSAGPLLQPPGLCSSLRPPL 1149
QY 139 DDYVNPESGESASASLDGSEYVNVSGELHGA 172
DB 1150 PDSLEGP-----GTLEGSPEIQAHVSDPVLPSS 1177

RESULT 7
US-10-193-692-2
Sequence 2, Application US/10193692
Publication No. US20030074682A1
GENERAL INFORMATION:
APPLICANT: Exelixis Deutschland GmbH
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF A NOVEL TELBOST POTASSIUM
FILE REFERENCE: ARO2-005C
CURRENT APPLICATION NUMBER: US/10/193,692
```

```

; CURRENT FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1186
; TYPE: PRT
; ORGANISM: Danio rerio
US-10-193-692-2

```

Query Match	10.5%;	Score 112.5;	DB 14;	Length 1186;
Best Local Similarity	31.6%;	Pred. No. 5.4;		
Matches 49;	Conservative 20;	Mismatches 61;	Indels 25;	Gaps 10;

Qy 27 VAPMPAYEVTSYEPLSQDILLPRPSQPL--GSHRPPSSRRSDSGANVSAYENEE 84
Db 1048 MAPPPAYEAV--SPDLAH--VP--PAHPPSLTYTAAHNTTPSLQITD-----ASSPGKS 10368

Dy
85 PACDADDEDDEDDYHNPQ-YLVLEPDST----PATSTAASA-PALSTPGIRDSAFSMSEI 138

Dd
1099 PDVDSLKEKSPDLSLSCGHHVTASTDTMSMSPETLTSLVPAGPLLOPGLCLSSLRFPRL 1156

QY 139 DDYNNVPESGESAEASLDGSRVY-NVSGELHPGA 172
| : | : | | : | : | : | :
Db 1159 PDSLEGP-----GTLESGPEIQRHVSDFVLPGS 1186

RESULT 8
US-10-73

; Sequence 9638, Application US/10739930
; Publication No. US20040216190A1

APPLICANT: KOVALIC, David K.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT

CURRENT APPLICATION NUMBER: US/10/739,930
 CURRENT FILING DATE: 2003-12-18
 NUMBER OF SEQ ID NOS: 11088

```

; LENGTH: 430
; TYPE: PRT
; ORGANISM: Triticum aestivum
;

```

```

; NAME/KEY: unsure
; LOCATION: (1)..(430)
; OTHER INFORMATION: unsure at all Xaa locations
;

```

OTHER INFORMATION: Clone ID: TRIAE-23APR03-C12877_1.f
US-10-739-930-9638

Query Match	10.5%;	Score 112;	DB 17;	Length 430;
Best Local Similarity	25.7%;	Pred. No. 1.7;		
Matches	64;	Conservative	26;	Mismatches 75;
				Indels 84;
				Gaps 14;

QY 21 FKRR-HTVAVMPRAYPPVTSVYPPLSQDPLPI PRS-----POLPGGSHRTP----- 65
 Db 21 FKGSHTVLGSGPPPSASTSSSR--SNPNPNPNRPSVPKQGPFRPAGPSEETPTFTAVISS 79

Db Oy 66 SSRR-DSDDGNSVASYENERPACEDADEBD-DYHNPGLVLVPDSTPTSTA-----117
80 SSRPDANGSDTVAC-----PNCGDASFSESHAVSEHLDG---CLASAGGARARAAAYLKG 131

```

OY      118  ---PSA-----PALSTPGIDSAFMSM-----SIDY  141
          :|
          :|
DB      132  NPPAAAVEVKKLLGNLKEPGRNDKRYRVRVLGNPRIKEAVADRREGGLLELLEAVGFTIGD-  190

```

OY | | | | | | 193
 |_ _ _ _ |_ _ _ _ |_ _ _ _ |_ _ _ _ |_ _ _ _ |_ _ _ _ |
142 UNPESGE-----SAEKSLDGRREYNVVSQELHPGAKTEPALSSQEAAEEVEEGGA

Db | | | | | | 243
 |_ _ _ _ |_ _ _ _ |_ _ _ _ |_ _ _ _ |_ _ _ _ |_ _ _ _ |
191 -----EGGLPAVMDETPEARLRIGRAVLLEHRSHPAPL---PTLAGSESKESCRCGV

QY	194 PDYENLQEL 202
	::: ::
Db	244 DEQKEVNKI 252

RESULT 9
US-10-113-794A-2

; Sequence 2, Application US/10113794A
; Publication No. US20030022202A1
; Patent Information

;; APPLICANT: Flanagan et al.
;; TITLE OF INVENTION: B-GLUCURONIDASE REGULATION OF G-PROTEIN COUPLED
;; TYPE OF INVENTION: CHEMOTHERAPY

```

;; FILE REFERENCE: 2535/106
;; CURRENT APPLICATION NUMBER: US/10/113,794A
;; CURRENT FILING DATE: 2002-04-01
NUMBER OF SEQ ID NOS: 6

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 519
; TYPE: DMT

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-113-794A-2

```

Query Match	10.5%	Score 112;	DB 14;	Length 519;
Best Local Similarity	24.3%;	Pred. No. 2.2;		
Matches	51;	Conservative	22;	Mismatches 79;
				Indels 58;
				Gaps 11.

```

QY      3 PGSDYDSTSSLSLYPRGIIQFKRHVAWPAPYAPVVISYPLSPQ-----DLPP-----IPRSP    53
Db      57 PSQEPLSSKDSATSEG-----SFGPRDAPPSKOVPRCGEPPPAQDLSPCQDLPAQ    10

```

[illegible]

Db **TGDP-----AAP-RPAFVPEVRLDSTYSQKA-----GAEGCGSGDEEDAEAA** **20**

```

QY      166 QELHPGAKTBPALSSQEAEEVEEBEGAP 195
          :|:|:|:|:|
Db      207 EEVEEG-----EEGDEDEDEDTSD 225

```

RESULT 10
US-10-428-487-14
: Sequence 14, A

;; sequence 14, Application US/1042848
;; Publication No. US20040006780A1
: GENERAL INFORMATION:

```

; APPLICANT: KASTIBELLI, LUCA K.
; APPLICANT: GERBER, HANS-PETER
; TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
; FILE REFERENCE: 09800080-0103

```

;; CURRENT APPLICATION NUMBER: 05/10/42
;; CURRENT FILING DATE: 2003-05-02
;; PRIOR APPLICATION NUMBER: 09/815,153
;; PRIOR FILING DATE: 2001-03-21

```

; PRIOR APPLICATION NUMBER: 60/191,201
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 14
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo sapiens

```

Query Match 10.5%; Score 112; DB 15; Length 519;
Best Local Similarity 24.3%; Pred. No. 2.2;
S-10-428-48/-14

```

QY      3 PGASYDSTSSDSLPRGCIQFKRPHIVAPWPAPYPPVTSYSPLSQP-----DLLP--IPRSP 55
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     57 PSQGEPLSKSDATSFG-----SPPGDAPFSPKDVPPQGEPPPAQDLSFCQDLPLAQ 107

```

```

APPLICANT: Zerhusen, Bryan D.
APPLICANT: Gorman, Linda
APPLICANT: Shenoy, Suresh G.
APPLICANT: Pena, Carol E.A.
APPLICANT: Smithson, Glenda
APPLICANT: Burgess, Catherine E.
APPLICANT: Gerlach, Valerie
APPLICANT: Padigaru, Muralidhara
APPLICANT: Shinkels, Richard A.
APPLICANT: Gangolli, Esna A.
APPLICANT: Taupier Jr., Raymond J.
APPLICANT: Caeman, Stacie J.
APPLICANT: Ji, Weizhen
APPLICANT: Anderson, David W.
APPLICANT: Liete, Mario W.
APPLICANT: Rastelli, Luca
APPLICANT: Edinger, Shlomit R.
APPLICANT: Stone, David J.
APPLICANT: MacDougall, John R.
APPLICANT: Rothenberg, Mark E.
TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Samea
FILE REFERENCE: 21402-322C
CURRENT APPLICATION NUMBER: US/10/114,270
CURRENT FILING DATE: 2002-11-27
PRIOR APPLICATION NUMBER: 60/281,086
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,136
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,663
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/281,906
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/282,020
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,930
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/282,934
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/283,512
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/283,710
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/284,234
PRIOR FILING DATE: 2001-04-17
Remaining Prior Application data removed - See file Wrapper or PALM.
SEQ ID NO 152
LENGTH: 923
TYPE: PRT
ORGANISM: Homo sapiens
US-10-114-270-152

Query Match      10.5%; Score 112; DB 15; Length 923;
Best Local Similarity 24.3%; Pred. No. 4.4;
Matches 51; Conservative 22; Mismatches 79; Indels 58; Gaps 11;

3 PGSDYSSSLVYRGIQFKRPHTVAMPAPVPPVTSYPLSQP---DILP---IPRSP 55
462 PSQGPLSSKKSATSEG-----SPGPDAPPSKDVPPCQBPAPQADLSFPCODLPAGQ 512
56 QPLGSGHRTSSRRDSDGANSVASYENEBACEDADE-----DDYHNPGLVLT 106
513 EPL--PHQDPLTRTDLPALQe--SPTRDLPCCQDLPSSQSLPAKALTEDTWSGGDLAA 568
107 PDSTPATSTAAPSPALSTGIR-DSAFSMESIDDIYVNEESGSAASALDSGREYVNV 165
569 TGDP-----AAP-RPAVPIEVRIIDSTYSOKA-----GAEGCGSGDEDAERA 611
166 QELHGAAKTEPALSSQEAEEVBEBCAPD 195
612 EEVEEG-----EEGEBDEDEDTSD 630

```

RESULT 13
US-10-425-115-224853
; Sequence 224853, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/425.115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 224853
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) (169)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_136658C.1.pep
US-10-425-115-224853

Query Match 10.4%; Score 111; DB 17; Length 169;
Best Local Similarity 28.8%; Pred. No. 0.67; Mismatches 60; Indels 24; Gaps 4;
Matches 40; Conservative 7;
Qy 1 RLPSYSTSDSYPRGIGFKRPHTVAPW-----PAYPVTSYPLSQPDL 48
Db 27 RNPXPRPTPRXALPRP-----PSTPRPRXALPRQAQPTXYPPL--QPRSPDR 79
Qy 49 LPIRSPQLGSHRTSPSSRDGANSVASYENEEPCADEDEDDYHNGYLVLDP 108
Db 80 PPPPLTSPCCPRRHPRAPSPPPRRMPA-----PPCVPRPRGRSPHPPPTDASP 134
Qy 109 STPATSTAPASAPALSTPG 127
Db 135 TPPKPLTPAPHTPALPTPG 153

RESULT 14
US-10-437-963-184670
; Sequence 184670, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437.963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 184670
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) (346)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81641C.1.pep

US-10-437-963-184670
Query Match 10.3%; Score 110; DB 16; Length 346;
Best Local Similarity 24.0%; Pred. No. 1.9; Mismatches 41; Conservative 22; Indels 78; Gaps 5;
Matches 41;
Qy 24 PHTVAPWPPAYPP-----VTSYPLSQPDLPIRSPQLGSHRTSPSSRRSDGA 74
Db 163 PSSPSP-PPASPPSPSSGAAPTTPPPASPPSTPTPASPAPSPAPPS-----A 215
Qy 75 NSVASYENEEPCADEDEDDYHNGYLVLDPSTPATSTAPASAPALSTPGIRDSAFS 134
Db 216 NAPSQAQARNPSATSPRSR-----STAPAPSSSAASTPTAPAPARSSSSWS 262
Qy 135 MESIDYVNVPESGESAASLDGSEYVNVSOELHPCGAATPAPALSSQGA 185
Db 263 PTXTGSSASPPSSSSPSSGGGGGGAAPAGAPAPAT-PAGTNSSPA 312

RESULT 15
US-10-156-761-12405
; Sequence 12405, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156.761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12405
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12405

Query Match 10.3%; Score 110; DB 14; Length 584;
Best Local Similarity 25.3%; Pred. No. 3.6; Mismatches 50; Conservative 24; Indels 82; Gaps 11;
Matches 50;
Qy 16 PRGIQFKRPHTVAPWPPAYPVTSYPLSQPDLPIRSPQLGSHRTSPSSRRSDGAN 75
Db 315 PTGAFGAPPTTATP-PPGPAP---WEPPAPPEPPAP-APAPAGRTPTDAGNIDSGS 369
Qy 76 SVASYENEEPCADEDEDDYHNGYLVLDPSTPATSTA-APSAPALSTPGIRDSAFS 134
Db 370 G-----SCADBDLPYAPAPW-----KTPAPRTVPAVEVEVEKPPVRE---- 409
Qy 135 MESIDYVNVPESGESAASLDG--SREYVNVSOELHPCGAATPAPALSSQGAEBVEEG 192
Db 410 FDSVAEAVLAPESGSGPDGAAGPLAPMARINEAVKG--RIEAAAMAEAA---VHG 463
Qy 193 A-----PDYENLOEL 202
Db 464 AATLGPHEPDVLRREL 481

Search completed: November 21, 2004, 14:32:23
Job time : 143 secs

Db 12 PGSRRSSIHSGPEELPTPEPDPF-----PPPLPMASSHNGGPEPPCPSPSEAMPLT 64

Qy 38 TSYPLSQDPLPIPTRSPOPLGSGHRTSSRRDSDGANSVASYENEPACEDADEDDDDY 97

Db 65 TSSAPGCEPALIPGHEPEPPGGS---PALLRDLAATPAS----- 103

Qy 98 HNPGLVLP-DSTPATs---TAAP-----SAPALSTPGIRDSAPsME----- 136

Db 104 ---PLVLLPETRAPEEPQPSAPSPHVKPLSAAPVEGSDRKOSNLSIALSSGLEKL 159

Qy 137 -----SIDDVNVVSEGESAE-----ASLDGSRHYVNVSOELHPGAKTEPALSSQ 183

Db 160 KTVTSGSIQPTVQAQAGQMVDTKRLKDSAVLDQSAKYHLTHD-----ELISLLQ 211

Qy 184 EAEVEEGAPDYENLQEL 202

Db 212 RERELSRD---EHVQEL 226

RESULT 4

138073 nucleolar phosphoprotein p130 - human

C:Species: Homo sapiens (man)

C:Date: 17-May-1996 #sequence _revision 17-May-1996 #text_change 09-Jul-2004

C:Accession: 138073; S52282

R:Pat, C.Y.; Chen, H.K.; Sheu, H.L.; Yeh, N.H.

J. Cell Sci. 108, 1911-1920, 1995

A:Title: Cell-cycle-dependent alterations of a highly phosphorylated nucleolar protein p130

A:Reference number: 138073; PMID:95386590; PMID:765714

A:Accession: 138073

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-699 <RES>

A:Cross-references: UNIPROT:Q14978; EMBL:Z34289; NID:9663007; PIDN:CAA64063.1; PID:9663006

C:Superfamily: nucleolar-cytoplasm shuttle phosphoprotein

C:Keywords: phosphoprotein

Query Match 11.3%; Score 121; DB 2; Length 699;

Best Local Similarity 27.3%; Pred. No. 0.65;

Matches 59; Conservative 23; Mismatches 70; Indels 64; Gaps 10;

Qy 5 SYDSTSSSLVRYGQIFKRPHTVAPWPPA-----YPPVTSYPLSQDPLPIPTSPQ 56

Db 362 SSDSDSDS-----SEDDAPESKPGAGTTKNSNKKPAVTKSPAVKPAAP----KQ 408

Qy 57 PLGSHRTPTSSRRDSDGANSVASYENEPACEDADEDDDYHNPGLVLPDSTPATSTA 116

Db 409 PVGGGQKLTLRADSSSSSEESSSSSEKTKK-----MVATTKKATAKA 453

Qy 117 APSAPALSTP-GIRDSAPMESIDYVNVPEGESASAPASLDGSRHYVNV-QELHPGAK 174

Db 454 ALSLPKAKAPQGSRRDS--SSDS-----DSSSSEEEKTSKSAVKKRPQVAGGAAP 503

Qy 175 TEPAL-----SQGAEEVEE-----GAP 194

Db 504 SKPASAKKGAESSNSSSDSDSSEEEKLKKGKSP 539

RESULT 5

S26058

probable transforming protein (can) - human

C:Species: Homo sapiens (man)

C:Date: 13-Jan-1995 #sequence _revision 13-Jan-1995 #text_change 09-Jul-2004

C:Accession: S26058

R:van Lindern, M.; Fornetord, M.; van Baal, S.; Jaegle, M.; de Wit, T.; Buijs, A.; Grosveld, G.

Mol. Cell. Biol. 12, 1687-1697, 1992

A:Title: The translocation (6;9), associated with a specific subtype of acute myeloid leukemia, can mRNA.

A:Reference number: S26058; PMID:92195315; PMID:1549122

A:Accession: S26058

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2090 <VON>

A:Cross-references: UNIPROT:P35658; EMBL:X64228; NID:G29652; PIDN:CAA45535.1; PID:G29653

Query Match 10.8%; Score 116; DB 2; Length 2000;

Best Local Similarity 24.9%; Pred. No. 5.3;

Matches 52; Conservative 20; Mismatches 85; Indels 52; Gaps 7;

QY 5 SYDSTSSDLSYPRGIGFQKRPHTVAPWPPVPTSTYPLSQPDLPIPRSPQ-PLGSGSR 63

DB 1440 SFGSGQGTSTVP-----PSAPPTTATPLTSPFTLSFGSLSSATTPLSPMNGRS 1492

QY 64 T-----PSSRRSDGANSVASYENEPACEDDEDDYHNGYLVLDPDSTPATST 115

DB 1493 TEEATSSALPEKPGSEVSAASAALIEEQGSAQ-----LQAPQPTSD 1535

QY 116 AAPSAPALSTPGIRG-----APSMESIDDVYVNPESGSAEASLDGSRVYNS 165

DB 1536 SYKKEPVLAAQPAVNSGTASSTSLVALSAEATPATGVDA--RTEAVPPASSFSV--- 1590

QY 166 QELHPGAATKEBPALSSQEAEEVEEGAP 194

DB 1591 ----FGQTRAVTAAIISAGPAVETISSTP 1615

RESULT 6

hypoetical protein 45 - equine herpesvirus 2

C:Species: equine herpesvirus 2

C>Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #ext_change 09-Jul-2004

C:Accession: S55640

R:Telford, B.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.

J. Mol. Biol. 249, 520-528, 1995

A:Title: The DNA sequence of equine herpesvirus 2.

A:Reference number: S55594; MUID:95302501; PMID:7783207

A:Accession: S55640

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-321 <TEL>

A:Cross-references: UNIPROT:066648; GB:U20824; NID:G695172; PIDN:AAC13833.1; PID:G695218

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995

Query Match 10.6%; Score 113.5; DB 2; Length 321;

Best Local Similarity 25.0%; Pred. No. 0.83;

Matches 47; Conservative 23; Mismatches 61; Indels 57; Gaps 8;

QY 31 PPAVPTVSYPLSGPDLPIPRSPQPLGSHRTSSRRSDGANSVASYENB----- 84

DB 78 PPSHVPVPSKPK-----PRT-----RAGCAADSDSESESDGDEDEETLHSQ 120

QY 85 ---PACEDADEDEDDYHNGYLVLDPSTPATSTAAPALSTPGIRDSAPSMESIDY 141

DB 121 DTPGSGSSSDSDDD-----QKLPATGCIKMPGYMS-RISDSSSSSSSSSD- 167

QY 142 VNVPSGSAEASALDSGR-----YVNVSGELHPGAATKEPA-----ALSQEA 186

DB 166 ---SSSSSSSDSESDGSTPEPDLIRQVTSGLAGVSPRAKPPAKKEVVISLSE 224

QY 187 EYEEGAP 194

DB 225 ESDSEGP 232

RESULT 7

hypoetical protein ZK84.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 09-Jul-2004

C:Accession: T29018

R:Kirsten, J.

submitted to the EMBL Data Library, April 1995

A:Description: The sequence of C. elegans cosmid ZK84.

A:Reference number: Z20553

A:Accession: T29018

A:Status: preliminary; translated from GB/EMBL/DBJ

Query Match 10.5%; Score 112; DB 2; Length 519;

Best Local Similarity 24.3%; Pred. No. 1.9;

Matches 51; Conservative 22; Mismatches 79; Indels 58; Gaps 11;

QY 3 PGSYDSTSSDLSYPRGIGFQKRPHTVAPWPPVPTSTYPLSQP-----DLIP---IPRP 55

DB 57 PSQEPPLSSKDSATSG-----SPGPDAPPSKDVPPCGPPADLSPQDLPAQ 107

QY 56 QPLGSHRTSSRRSDGANSVASYENEPACEDDEDE-----DDYHNGYLVL 106

DB 108 EPL--PHQDPLITKLPALIG--SETRDLPPGQDLPSPQVSLPAKALREDTMSSGDLAA 163

QY 107 PDSTPATSTAAPALSTPGIR-DSAPSMESIDYVNVPSGSAEASLDGSRVYNS 165

DB 164 TGDPP-----AAP-RAFYIPEVRLDSTYSOKA-----GAEQCGSGDEEDAEBA 206

QY 166 QELHPGAATKEBPALSSQEAEEVEEGAP 195

A:Molecule type: DNA

A:Residues: 1-801 <KIR>

A:Cross-references: UNIPROT:Q23635; EMBL:U23181; PIDN:AAC48204.1; GSPDB:GN00020; CESP:Z

A:Experimental source: strain Bristol N2; clone ZK84

C:Gene: CESP:ZK84.1

A:Map position: 2

A:Introns: 22/2; 45/3; 108/1

C:superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology

Query Match 10.6%; Score 113.5; DB 2; Length 801;

Best Local Similarity 24.6%; Pred. No. 2.5;

Matches 58; Conservative 29; Mismatches 78; Indels 71; Gaps 11;

QY 4 GSYDSTSSDLSYPRGIGFQKRPHTVAPWPPA-----YPPYTSYPLSQ-----PDLIP- 50

DB 157 GSGDAPASNSVTETLTATPAIILP-PPAEPRKIIIPASGVYTSQEQAGADAPAPAV 215

QY 51 IPRSPQPLGSHRTSSRRD-----DG-----ANSVASYENEPACEDDEDD 95

DB 216 VETPAPTPAETPAPATEASNAVGTPEGYVDGTAANAASAVAPAVETPAPAPAE-- 273

QY 96 DYHNGYLVLDPSTPATSTAAPALSS-----TPGIRDSAPSMESIDYVNVPSGES 150

DB 274 -----TPAPATSEEAAPAPAPAEPTAPPTVGAAPAAANSYSA--CGDA 317

QY 151 AEAASLDGSRVYVNVSGELHPGAATKEPPA-----LSQEAEEVEEGAP 195

DB 318 ATAPAP-----SSEADAAAPDTSASADTTLAALVDTSSSEHAETEAATD 365

RESULT 8

G-protein signaling regulator RGS3 - human

C:Species: Homo sapiens (man)

C>Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #ext_change 09-Jul-2004

C:Accession: S78089

R:Druey, K.

submitted to the EMBL Data Library, May 1995

A:Reference number: S78089

A:Accession: S78089

A:Molecule type: mRNA

A:Residues: 1-519 <DRU>

A:Cross-references: UNIPROT:49796; EMBL:U27655; NID:G1216368; PID:G1216369

R:Druey, K.M.; Blumer, K.J.; Kang, V.H.; Kehrl, J.H.

Nature 379, 742-746, 1996

A:Title: Inhibition of G-protein-mediated MAP kinase activation by a new mammalian gene

A:Reference number: A56012; MUID:96178495; PMID:8602223

A:Accession: S68436

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-355, 'K', 357-519 <DRW>

A:Cross-references: EMBL:U27655

A:Experimental source: tonsil

A>Note: the sequence from Fig. 1b is inconsistent with that from Fig. 1a in having 356-7

Query Match 10.5%; Score 112; DB 2; Length 519;

Best Local Similarity 24.3%; Pred. No. 1.9;

Matches 51; Conservative 22; Mismatches 79; Indels 58; Gaps 11;

QY 3 PGSYDSTSSDLSYPRGIGFQKRPHTVAPWPPVPTSTYPLSQP-----DLIP---IPRP 55

DB 57 PSQEPPLSSKDSATSG-----SPGPDAPPSKDVPPCGPPADLSPQDLPAQ 107

QY 56 QPLGSHRTSSRRSDGANSVASYENEPACEDDEDE-----DDYHNGYLVL 106

DB 108 EPL--PHQDPLITKLPALIG--SETRDLPPGQDLPSPQVSLPAKALREDTMSSGDLAA 163

QY 107 PDSTPATSTAAPALSTPGIR-DSAPSMESIDYVNVPSGSAEASLDGSRVYNS 165

DB 164 TGDPP-----AAP-RAFYIPEVRLDSTYSOKA-----GAEQCGSGDEEDAEBA 206

QY 166 QELHPGAATKEBPALSSQEAEEVEEGAP 195

Db 207 EHVREG-----EEGEDEDETRSD 225

RESULT 9

B40505 hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhouser or Becker)

C/Species: suid herpesvirus 1
C/Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 09-Jul-2004
C/Accession: B40505

R/Chenung, A.K.

J. Virol. 65, 5260-5271, 1991

A/Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies virus.

A/Reference number: A40505; MUID:91374576; PMID:1654441

A/Accession: B40505

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-1958 <CHE>

A/Cross-references: UNIPROT:Q69340; GB:M57505; NID:G334066; PIDN:AAA47468.1; PID:G334068
C/Superfamily: pseudorabies virus 1 nuclear antigen

Query Match

Best Local Similarity 10.5%; Score 112; DB 2; Length 1958;
Best Local Similarity 35.3%; Pred. No. 9.2;

Matches 30; Conservative 5; Mismatches 26; Indels 24; Gaps 3;

29 PMPAYPVTVSYPLSQPDLPIRSPQPLGGS-----HRTSSRRDS 71

483 PSEPPRP-----PPLPPPPPPPPPPPPAGGSAARRRRGGGGGPPRGGRKRRRA 538

72 DGANSVASYENEPACEDADEDD 96

539 EGTEAAADAE-----EDGEDDE 560

RESULT 10

136729

probable serine/threonine-specific protein kinase (EC 2.7.1.-) - Streptomyces coelicolor

C/Species: Streptomyces coelicolor

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Aug-2004

C/Accession: T36729

R/Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1999

A/Reference number: Z21612

A/Accession: T36729

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-576 <MUR>

A/Cross-references: UNIPROT:Q9XA04; EMBL:AL079308; PIDN:CAB45227.1; GSPDB:GN00070; SCOE

A/Experimental source: strain A3(2)

C/Genetics:

A/Gene: SCOE:DB:SCHE9.30

C/Superfamily: protein kinase homology

C/Keywords: phosphotransferase; serine/threonine-specific protein kinase

Query Match

Best Local Similarity 10.4%; Score 111.5; DB 2; Length 576;
Best Local Similarity 25.5%; Pred. No. 2.3;

Matches 49; Conservative 12; Mismatches 74; Indels 57; Gaps 7;

7 DSTSSDSLPRGICFKRPHVAPMPAYPVTVSYPLSQPDLPIRSP-----OPLGGS 61

300 DDVPTDPRPAG-----RPPGVDPSPPATPP-----PPAPPMQGTTPPAGSPSGLDPRSPGS 352

62 HRTSSRRDSGANSVASYENEPACEDADEDDYHNPGLVVLDPSTPATSTAAAPAP 121

353 PGPPPTGDPSTPAP-----PPPTPTATGTTPAP 382

122 ALSTPGIRDSAFMSISIDYVNPESGESAEASLDGSRVYVNGQLPFGAAXTEPALIS 181

383 GL--PPASDGWT--PSTSGPTAPPSAPASAPSA-----PGTTRAPPHOTH 427

182 SQAEVEVEEGA 193

428 SEEVPLAERPGA 439

RESULT 11

S74716 hypothetical protein slr0962 - Synecocystis sp. (strain PCC 6803)

C/Species: Synecocystis sp.

A/Variety: PCC 6803

C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C/Accession: S74716

R/Kaneko, T.; Sato, S.; Kori, H.; Tanaka, A.; Asami, E.; Nakamura, Y.; Miyajima, N.;

O., K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,

DNA Res. 3, 109-116, 1996

A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis

S.

A/Reference number: S74322; MUID:97061201; PMID:8905231

A/Accession: S74716

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-477 <KAN>

A/Cross-references: UNIPROT:P72852; EMBL:D90901; GB:AB001339; NID:G1651897; PIDN:BAA16867

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match

Best Local Similarity 10.4%; Score 111; DB 2; Length 477;
Best Local Similarity 24.9%; Pred. No. 2;

Matches 48; Conservative 24; Mismatches 89; Indels 32; Gaps 5;

33 AYPPVTSYPLSQPDLPIRSPQPLGSHRTSSRRDSGANSVASYENEPACEDADE 92

256 AVPPV--QPKASQPEVEQKVAIEPESSTACTDDMDDEGDEGDEPSTEDSSPGEISDE 313

93 DEDYHNP-----YLVLPDSTPATSTAAAPALSTPGIRDSAF 133

314 VSED--NPENDQKVAIVETVQIEQITLIEVDSNITAEAPAPALIEEQIAGE-- 369

134 SMESIDYVNPESGESAEASLDGSRVYVNGQLPFGAAXTEPALISQEA-----EV 188

370 --ELVVEVAVIAPMSGITTAIVETVPELVPEFTNIEGQPSGVDEYERETETETGEI 427

189 EEEGAPDYENQ 201

428 TEATASDLEEVPE 440

RESULT 12

T20884

hypothetical protein F14D7.2 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T20884

R/Berks, M.

submitted to the EMBL Data Library, July 1996

A/Reference number: Z19340

A/Accession: T20884

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-457 <WIL>

A/Cross-references: UNIPROT:Q19452; EMBL:Z77658; PIDN:CAB01156.1; GSPDB:GN00023; CESP:Fi

A/Experimental source: clone F14D7

C/Genetics:

A/Gene: CESP:F14D7.2

A/Map position: 5

A/Intons: 31/2; 257/1; 287/3; 320/3; 343/3; 380/3

Query Match

Best Local Similarity 10.3%; Score 110; DB 2; Length 457;
Best Local Similarity 20.0%; Pred. No. 2.2;

Matches 57; Conservative 32; Mismatches 80; Indels 116; Gaps 10;

1 RLP--GSYDSTSSDSLPRGICF-----KRPHTVAPMP-----PAYPV 37

28 RLPSGSENGTSGSFFRGMNSNSESADDDKVIKQIPETSSSPHRSRRKAPPS 87

38 TSYPLSQPDL-----PIRSP 55

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 14:10:14 ; Search time 194 Seconds

(without alignments)
602.067 Million cell updates/sec

Title: US-09-597-920b-4_COPY_31_233

Perfect score: 1071

Sequence: 1 RLPGSYDSTSSDLYPRGIQ.....BAEVEBEGAPDYENQLQLN 203

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1046.5	97.7	262	1	LAT_HUMAN
2	683	63.8	242	1	LAT_MOUSE
3	683	63.8	242	2	CAD27390
4	678	63.3	242	2	Q80UR6
5	670.5	62.6	241	1	LAT_RAT
6	665	62.1	242	2	Q8C2E5
7	132	12.3	710	2	Q9CAL8
8	128	12.0	1296	2	Q6BR88
9	126.5	11.8	754	2	Q6PEI3
10	126.5	11.8	754	2	AAH58052
11	123.5	11.5	311	2	Q6L4T3
12	123.5	11.5	311	2	AAT38083
13	122	11.4	704	1	NP14_RAT
14	121.5	11.3	253	2	Q9UPM0
15	121.5	11.3	367	2	Q9NEK0
16	121	11.3	418	2	Q96C17
17	121	11.3	699	1	NP14_HUMAN
18	121	11.3	700	2	Q9BHV3
19	119	11.1	703	2	Q6ZOK6
20	119	11.1	703	2	BAC37850
21	118.5	11.1	700	2	Q8CE21
22	118.5	11.1	1027	2	Q7Q2X2
23	118	11.0	686	2	Q874W2
24	116	10.8	1519	2	Q7Z3C4
25	116	10.8	2090	1	NP14_HUMAN
26	116	10.8	2091	2	Q75R47
27	116	10.8	2091	2	BAD07398
28	116	10.8	2093	2	Q81X10
29	115	10.7	490	2	Q968P7
30	115	10.7	490	2	Q968P7
31	115	10.7	490	2	Q968S0

ALIGNMENTS

RESULT 1	LAT_HUMAN	STANDARD	PRT	262 AA.
AC	Q43561; Q43919;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	05-UTR-2004 (Rel. 44, Last annotation update)			
DE	Linker for activation of T cells (36 kDa phospho-tyrosine adaptor protein) (pp36) (p36-38).			
GN	Name=LAT;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	(1)			
RP	SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT), AND SEQUENCE OF 32-47			
RP	AND 219-233.			
RC	TISSUE=Leukemia;			
RX	MEDLINE=98149346; PubMed=9489702;			
RT	Zhang W., Sloan-Lancaster J., Kitchen J., Tribble R.P., Samelson L.E.;			
RT	"LAT: the ZAP-70 tyrosine kinase substrate that links T cell receptor			
RT	to cellular activation.";			
RT	Cell 92:83-92(1998).			
RL	(2)			
RP	SEQUENCE FROM N.A. (ISOFORM SHORT).			
RC	TISSUE=Thymus;			
RX	MEDLINE=98197173; PubMed=9529333;			
RT	Weber J.R., Orstavik S., Torgersen K.M., Danbolt N.C., Berg S.F.;			
RT	Ryan J.C., Taeken K., Imboden J.B., Vaage J.T.;			
RT	"Molecular cloning of the cDNA encoding pp36, a tyrosine-			
RT	phosphorylated adaptor protein selectively expressed by T cells and			
RT	natural killer cells.";			
RT	J. Exp. Med. 187:1157-1161(1998).			
RL	(3)			
RP	SEQUENCE FROM N.A. (ISOFORM SHORT).			
RC	TISSUE=Colon;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RT	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.;			
RT	Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.;			
RT	Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.;			
RT	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.;			
RT	Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.;			
RT	Stapleton M., Soares M.B., Bonaldi M.F., Caesvant T.L., Schetz T.E.;			
RT	Brownstein M.D., Usdin T.B., Toshiyuki S., Carninci P., Prange C.;			
RT	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.;			
RT	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.;			
RT	Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.;			
RT	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.;			
RT	Faney J., Helton E., Kettelman M., Madden A., Rodriguez S., Sanchez A.;			
RT	Blakesley R.W., Tschman J.W., Green E.D., Dickson M.C.;			
RT	Rodriguez A.C., Grouse L.H., Schmitt J., Myers R.M.;			
RT	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.;			
RT	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;			

32	115	10.7	1146	2	Q86XA7	Q86XA7 homo sapien
33	114.5	10.7	838	2	Q9VOA9	Q9VOA9 drosophila
34	114.5	10.7	917	2	Q8IUQ1	Q8IUQ1 homo sapien
35	114.5	10.7	1203	2	Q9ACX5	Q9ACX5 streptomyc
36	114	10.6	490	2	Q968P6	Q968P6 plasmodium
37	114	10.6	2183	2	Q6CEV2	Q6CEV2 yarrowia 11
38	113.5	10.6	321	2	Q66648	Q66648 equid herpe
39	113.5	10.6	801	2	Q23635	Q23635 caenorhabd
40	113.5	10.6	1101	2	Q6BK14	Q6BK14 debaromyce
41	113.5	10.6	1374	2	Q9V5U0	Q9V5U0 drosophila
42	113.5	10.6	1449	2	Q90112	Q90112 drosophila
43	113.5	10.6	1450	2	Q8IQB8	Q8IQB8 drosophila
44	113.5	10.6	2409	2	Q960G6	Q960G6 drosophila
45	113.5	10.6	2786	2	Q9VSU2	Q9VSU2 drosophila

RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Involved in T-cell antigen receptor (TCR) signal
 transduction pathway. May play an important role downstream of the
 activation of protein tyrosine kinases (PTKs).
 CC -1- SUBUNIT: When phosphorylated, interacts directly with the p85
 subunit of phosphoinositide 3-kinase and the SH2 domain of GRB2,
 GRAP, and PLC-gamma-1. Interacts indirectly with CBL, SOS, VAV,
 and SLP-76.
 CC -1- SUBCELLULAR LOCATION: Type III membrane protein (Potential).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named Isoforms=2;
 CC Name=long;
 CC IsoId=043561-1; Sequence=Displayed;
 CC Name=short;
 CC IsoId=043561-2; Sequence=VSP 004303;
 CC -1- TISSUE SPECIFICITY: Expressed in thymus, T cells, nk cells, mast
 cells and, at lower levels, in spleen.
 CC -1- PTM: Phosphorylated on tyrosines by ZAP-70 and SYK.
 CC Phosphorylation leads to the recruitment of multiple signaling
 molecules. is one of the most prominently tyrosine-phosphorylated
 proteins detected following TCR engagement.
 CC -1- MISCELLANEOUS: Engagement of killer inhibitory receptors (KIR)
 disrupts the interaction of PLC-gamma-1 with LAT and blocks target
 cell-induced activation of PLC, maybe by inducing the
 dephosphorylation of LAT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF036906; AAC39637.1; -
 DR EMBL: AF036905; AAC39636.1; -
 DR EMBL: AJ232280; CA11218.1; -
 DR EMBL: BC011563; AAH11563.1; -
 DR Genew; HGNC:18974; LAT.
 DR MIM; 602354; -
 DR GO; GO:0005070; F:SH3/SH2 adaptor protein activity; TAS.
 DR GO; GO:0006955; P:immune response; TAS.
 DR InterPro: IPR008359; LAT.
 DR PRINTS; PR01781; LATPROTEIN.
 KW Alternative splicing; Direct protein sequencing; Phosphorylation;
 KW Signal-anchor; Transmembrane.
 FT DOMAIN 1 4
 FT TRANSMEM 5 27
 FT 28 262
 FT DOMAIN 164
 FT 164 164
 FT DOMAIN 200 203
 FT 200 203
 FT DOMAIN 220 223
 FT 220 223
 FT MOD_RES 110 110
 FT 110 110
 FT MOD_RES 156 156
 FT 156 156
 FT MOD_RES 200 200
 FT 200 200
 FT MOD_RES 220 220
 FT 220 220
 FT MOD_RES 255 255
 FT 255 255
 FT VARSPPLIC 114 142
 FT 114 142
 FT SEQUENCE 262 AA; 27929 MW; BCD80AE7DCA64153 CAC64;
 Query Match 97.7%; Score 1046.5; DB 1; Length 262;
 Best Local Similarity 87.5%; Pred. No. 4.9e-58;
 Matches 203; Conservative 0; Mismatches 0; Indels 29; Gaps 1;

QY 1 RLPGSYDSTSDSLYPKGIQKRPHTVAPWPPAPVPTSYVPLSQDPLRIPSPQPLGG 60
 DQ 31 RLPGSYDSTSDSLYPKGIQKRPHTVAPWPPAPVPTSYVPLSQDPLRIPSPQPLGG 90
 QY 61 SHRTPSRRRSDGANSYASVENE-----EACGDAD 91
 DB 91 SHRTPSRRRSDGANSYASVENEASGIRGAQGWGWSWTRLTPVSLPEPEACDAD 150
 QY 92 EDEDYNNPGVLVLPSTPATSTPAAPSAPALSTPGIRDSAFSMESIDYVYNVPESESA 151
 DB 151 EDEDYNNPGVLVLPSTPATSTPAAPSAPALSTPGIRDSAFSMESIDYVYNVPESESA 210
 QY 152 EASLDGSEYVYNVQELHPGAKTEPPALSSQEAEEVEEGGAPDYENLQELN 203
 DB 211 EASLDGSEYVYNVQELHPGAKTEPPALSSQEAEEVEEGGAPDYENLQELN 262
 RESULT 2
 LAT_MOUSE STANDARD; PRT; 242 AA.
 AC 054557;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUN-2004 (Rel. 44, Last annotation update)
 DE Linker for activation of T cells (36 kDa phospho-tyrosine adaptor
 protein) (p36) (p36-38).
 GN Name:Lat;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBI_TaxID=10090;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thymus;
 RX MEDLINE=36149346; PubMed=9489702;
 RA Zhang W., Sloan-Lancaster J., Kitchen J., Tridle R.P., Samelson L.E.;
 RT "LAT: the ZAP-70 tyrosine kinase substrate that links T cell receptor
 to cellular activation.";
 RL Cell 92:83-92(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Salivary gland;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshlyuk S., Carninci P., Mullaly S.J.,
 RA Raha S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wotley K.C., Hale S., Garcia A.M., Gay L.U., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez P., Prange C.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Involved in T-cell antigen receptor (TCR) signal
 transduction pathway. May play an important role downstream of the
 activation of protein tyrosine kinases (PTKs) (By similarity).
 CC -1- SUBUNIT: When phosphorylated, interacts directly with the p85
 subunit of phosphoinositide 3-kinase and the SH2 domain of GRB2,
 GRAP, and PLC-gamma-1. Associates indirectly with CBL, SOS, VAV,
 and SLP-76 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type III membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: Thymus.
 CC -1- PTM: Phosphorylated on tyrosines by ZAP-70 and SYK.

Phosphorylation leads to the recruitment of multiple signaling molecules. Is one of the most prominently tyrosine-phosphorylated proteins detected following TCR engagement (by similarity).

- MISCELLANEOUS: Engagement of Killer Inhibitory receptors (KIR) disrupts the interaction of PLC-gamma-1 with LAT and blocks target cell-induced activation of PLC, maybe by inducing the dephosphorylation of LAT (by similarity).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF036907; AAC40054.1; -
DR MGD; MGI:1342293; Lat.
DR InterPro; IPR008359; LAT.
DR PRINTS; PR01781; LATPROTEIN.
KW Phosphorylation; Signal-anchor; Transmembrane.
FT DOMAIN 1 4 Extracellular (Potential).
FT TRANSMEM 5 28 Signal-anchor for type III membrane protein (Potential).
FT FT Cytoplasmic (Potential).
FT DOMAIN 29 242 PLC-GAMMA-1-BINDING MOTIF.
FT DOMAIN 136 178 GNB2-BINDING MOTIF 1.
FT DOMAIN 175 198 GNB2-BINDING MOTIF 2.
FT MOD_RES 113 113 Phosphotyrosine (by Tyr-kinases) (Potential).
FT MOD_RES 132 132 Phosphotyrosine (by Tyr-kinases) (Potential).
FT MOD_RES 175 175 Phosphotyrosine (by Tyr-kinases) (Potential).
FT MOD_RES 195 195 Phosphotyrosine (by Tyr-kinases) (Potential).
FT MOD_RES 235 235 Phosphotyrosine (by Tyr-kinases) (Potential).
SQ SEQUENCE 242 AA; 26014 MW; 6AC25F7ABE61A5C1 CRC64;

Query Match 63.8%; Score 683; DB 1; Length 242;
Best Local Similarity 68.9%; Pred. No. 2.6e-35;
Matches 146; Conservative 16; Mismatches 38; Indels 12; Gaps 6;

QY 2 LFGSYDSTSSDGLYPRGIQFKRPHTVAPMPA--YPPVTSTYPLSQPDLPIPRSPQPLG 59
DB 33 LPVSYDSTSTESLYPRSLIKPPQITVPTPAVSYPVLVTSFPPLKQPDLLPIPRSPQPLG 92
QY 60 GSHRTPSSRRDSDGANSVASYENEPACE--DADEDEDYHNPGLVLPDSTPATSTAA 117
DB 93 GSHRTPSSQQNDSDANSVASYENQEPACKNVADDEDDEDPN-GYLVLVLPDSSPAVPV 151
QY 118 PSAPALSTPGIRDSAPSMESIDYVNVPSGESAASLDGSEYVNVSOELHPGAKTTP 177
DB 152 SSAPVPSNPDLGDSAFVSESCEDYVNVPSSESASLSLDGSEYVNVSPQOP-VTRAE 210
QY 178 AALSQGEAE-EVEERG-----APDYENIQELN 203
DB 211 ASVNSQVEBDEGEEGVDGEAPDYENIQELN 242

RESULT 3
CAD27390 PRELIMINARY; PRT; 242 AA.
ID CAD27390; CAD27390;
AC CAD27390; PRELIMINARY; PRT; 242 AA.
DT 02-MAR-2004 (TRENBLrel. 27, Created)
DT 02-MAR-2004 (TRENBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TRENBLrel. 27, Last annotation update)
DE Linker protein.
GN LAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;
RN [1] -
RP SEQUENCE FROM N.A.
RC STRAIN=129/Ole;
RA Aguado B., Nunez-Cruz S., Miazek A., Richeine S., Mura A.M., Richeine M., Sainy D., He H.T., Malsen B., Malsen M.;
RT "Spontaneous and exaggerated T helper type 2 immunity in mice with a point mutation in LAT.";
RT Science 296:2036-2040(2002).
DR EMBL; AJ438435; CAD27390.1; -
FT CHAIN 1 242 POTENTIAL.
SQ SEQUENCE 242 AA; 26014 MW; 6AC25F7ABE61A5C1 CRC64;

Query Match 63.8%; Score 683; DB 2; Length 242;
Best Local Similarity 68.9%; Pred. No. 2.6e-35;
Matches 146; Conservative 16; Mismatches 38; Indels 12; Gaps 6;

QY 2 LFGSYDSTSSDGLYPRGIQFKRPHTVAPMPA--YPPVTSTYPLSQPDLPIPRSPQPLG 59
DB 33 LPVSYDSTSTESLYPRSLIKPPQITVPTPAVSYPVLVTSFPPLKQPDLLPIPRSPQPLG 92
QY 60 GSHRTPSSRRDSDGANSVASYENEPACE--DADEDEDYHNPGLVLPDSTPATSTAA 117
DB 93 GSHRTPSSQQNDSDANSVASYENQEPACKNVADDEDDEDPN-GYLVLVLPDSSPAVPV 151
QY 118 PSAPALSTPGIRDSAPSMESIDYVNVPSGESAASLDGSEYVNVSOELHPGAKTTP 177
DB 152 SSAPVPSNPDLGDSAFVSESCEDYVNVPSSESASLSLDGSEYVNVSPQOP-VTRAE 210
QY 178 AALSQGEAE-EVEERG-----APDYENIQELN 203
DB 211 ASVNSQVEBDEGEEGVDGEAPDYENIQELN 242

RESULT 4
Q80UR6 PRELIMINARY; PRT; 242 AA.
ID Q80UR6; Q80UR6;
AC Q80UR6; Q80UR6;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Linker for activation of T cells.
GN Name=Lat;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1] -
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6NCR; TISSUE=Hematopoietic Stem Cell;
RC MEDLINE=22388257; PubMed=1477932;
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G., Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.B., Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C., Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S., Krzywnicki M.I., Skalska U., Smalios D.E., Scherch A., Schein J.E., Jones S.J., Maira W.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]

RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6NCR; TISSUE=Hematopoietic Stem Cell;
RA	Strasbourg R.; (Mar-2003) to the EMBL/genbank/DBJ databases.
RL	Submitted (MAR-2003) to the EMBL/genbank/DBJ databases.
DR	EMBL; BC052340; AAH52340.1; -
DR	MCD; MGI:1342293; Lat.
DR	GO; GO:0016020; C:membrane; IEA.
DR	GO; GO:0007165; P:signal transduction; IEA.
DR	InterPro; IPR008359; LAT.
DR	PRINTS; PR01781; LATPROTEIN.
SQ	SEQUENCE 242 AA; 25984 MW; 966955DASCFAF61 CRC64;
Query Match	63.3%; Score 678; DB 2; Length 242;
Best Local Similarity	68.4%; Pred. No. 5.4e-35;
Matches 145; Conservative 16; Mismatches 39; Indels 12; Gaps 6	
Dy	2 LPSYDSTRSSDSLYRGYIQFKRHTVAWPPA--YPVTSYPLSQPDLIPRSPQLG 59
Dd	33 LPVSYSASTESTELYSRLIKPKPTTVTRTPAVSYPLVTSPFPRLQPDLLPIPRSPQLG 92
Dy	60 GSHRRPSRRPSDDGANSVASYNENEPAC--DADDEDYNNRGLVYLPOSTPATSTRA 117
Dd	93 GSHRRPSSQQNSDDANSVASYNQEPACKNVADDEDEDYPR-GYLVLPPSSPAVPV 151
Dy	118 PSAPALSPFGIRDSAFSMESIDDYNNVPESGSASALDGSEYNNVCOLHPGAKTPE 177
Dd	152 SSAPLPSPNDLCDSAFSVESCEDDYNNVESESAPSALSDGSREYNNVPEQP-VTRAEL 210
Dy	178 AALSQGEAE-EVEEGE-----APPDENIQLN 203
Dd	211 ASVNGQEVEDEGECEGVDCGEAPDYENIQLN 242
RESULT 5	
ID	LAT RAT STANDARD; PRT; 241 AA.
LAT RAT	
AC	070601;
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DT	05-JUL-2004 (Rel. 44, Last annotation update)
DE	Linker for activation of T cells (36 kDa phospho-tyrosine adaptor protein) (pp36) (p36-38).
DE	Name:Lat;
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX	NCBI_Taxid=10116;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=PVG; TISSUE=Natural killer cells;
RX	MEDLINE=98197173. PubMed=9529333;
RA	Weber J.C., Orskovik S., Torgersen K.M., Danbolt N.C., Berg S.F.,
RA	Ryen J.C., Tasken K., Imboden J.B., Vaage J.T.;
RT	"Molecular cloning of the cDNA encoding pp36, a tyrosine-
RT	phosphorylated adaptor protein selectively expressed by T cells and
RT	natural killer cells."
J. Exp. Med.	187:1157-1161(1998).
-1- FUNCTION:	Involved in T-cell antigen receptor (TCR) signal transduction pathway. May play an important role downstream of the activation of protein tyrosine kinases (PTKs) (By similarity).
-1- SUBUNIT:	When phosphorylated, interacts directly with the p85 subunit of phosphoinositide 3-kinase and the SH2 domain of GRB2, GRAP, and PLC-gamma-1. Associates indirectly with CBL, SOS, VAV,
-1- SUBCELLULAR LOCATION:	Type III membrane protein (Potential).
-1- TISSUE SPECIFICITY:	NK cells.
-1- PTM:	Phosphorylated on tyrosines by ZAP-70 and SYK.
-1- PHOSPHORYLATION LEADS TO THE RECRUITMENT OF MULTIPLE SIGNALING MOLECULES.	Is one of the most prominently tyrosine-phosphorylated proteins detected following TCR engagement (By similarity).
-1- MISCELLANEOUS:	Engagement of killer inhibitory receptors (KIR) disrupts the interaction of PLC-gamma-1 with LAT and blocks target cell-induced activation of PLC, maybe by inducing the

CC					dephosphorylation of LAT (by similarity).
CC					-----
CC					This SWISS-PROT entry is copyright. It is produced through a collaboration
CC					between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC					the European Bioinformatics Institute. There are no restrictions on its
CC					use by non-profit institutions as long as its content is in no way
CC					modified and this statement is not removed. Usage by and for commercial
CC					entities requires a license agreement (See http://www.ebi.ac.uk/announcement/
CC					or send an email to license@isb-eib.ch).
CC					-----
DR	EMBL; AJ001184;	CMA04577.1;	-.		
DR	RGD; 620802;	Lac.			
DR	InterPro; IPR008359;	LAT.			
DR	PRINTS; PR01781;	LATPROTEIN.			
KW	Phosphorylation;	Signal-anchor;	Transmembrane.		
FT	DOMAIN	1	4		Extracellular (Potential).
FT	TRANSMEM	5	28		Signal-anchor for type III membrane protein (Potential).
FT					Cytoplasmic (Potential).
FT	DOMAIN	29	241		Poly-Glu.
FT	DOMAIN	217	224		PLC-GAMMA-1-BINDING MOTIF.
FT	DOMAIN	136	139		GRR2-BINDING MOTIF 1.
FT	DOMAIN	175	178		GRR2-BINDING MOTIF 2.
FT	DOMAIN	195	198		phosphotyrosine (by Tyr-Kinases)
FT	MOD_RES	113	113		(Potential).
FT					phosphotyrosine (by Tyr-Kinases)
FT	MOD_RES	132	132		(Potential).
FT					phosphotyrosine (by Tyr-kinases)
FT	MOD_RES	175	175		(Potential).
FT					phosphotyrosine (by Tyr-kinases)
FT	MOD_RES	195	195		(Potential).
FT					phosphotyrosine (by Tyr-kinases)
FT	MOD_RBS	234	234		(Potential).
FT					phosphotyrosine (by Tyr-kinases)
SQ	SEQUENCE	241 AA;	26192 MW;	FEDSC4DBBCC4275 CRC64;	

Query March	62.6%	Score 670.5	DB 1	Length 241
Best Local Similarity	67.3%	Pred. No. 1.6e-34		
Matches 142	Conservative 16	Mismatches 42	Indels 11	Gaps 5

Qy	2	LPGSDYSTDSDSLYPRGIQFRKPTVAWPBPAA--YPEVTSYTPSLISQDPLLIPRSPQLG	59
Db	33	LPASVDSASTESTLYPRSLILKXPQITVPBPRPATSPPLVTSPPPLRQDPLLIPRSPQLG	92
Qy	60	GSHTTPSRRRSDANSVASYENEKPEAC--DAEDDDDDYNNPCYLVLVLPSTPATSTAA	117
Db	93	GSRRMPSSRRQSDANSVASTENQEPARKVDEDEDQDYPE-GYLVLVLPDSSPAAVPV	151
Qy	118	PSAPALSTPGIRDSAFMESIDIDVYVNPESGESASALDGSREYVYVNSQELHPGAKTPE	177
Db	152	SSAIVNPSPDLGDGSAFMSWECEDVYVNPESGESASALDGSREYVYVNSQDQAP-VIRAEI	210
Qy	178	AALSSQEAEEVEEFG-----APDYENQELN	203
Db	211	ASVTSQEEVEDEEEDVDGEAPDYENQELN	241

RESULT 6	
Q8C2E5	PRELIMINARY; PRT; 242 AA.
AC	Q8C2E5;
DT	01-MAR-2003 (TREMBLrel. 23, Created)
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE	Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:R330025108 product:linker for activation of T cells, full insert sequence.
DE	Name=Lat;
GN	Mus musculus (Mouse).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Scuriongnathi; Muridae; Murinae; Mus.
OX	NCBI_TaxId=10090;
KN	[1]
RP	SEQUENCE FROM N.A.

```

RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RA The PANTOM Consortium;
RT "The RIKEN Genome Exploration Research Group Phase I & II Team;
RL Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtration of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Katsunai T., Taahiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Matshiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multiplex capillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX STRAIN=NOD; TISSUE=Thymus;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirao T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kaubukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Kono H., Koyama M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RT Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.
RL EMBL; AK088763; BAC40554.1; -.
DR MGD; MGI:1342293; Lat.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR008359; LAT.
DR PRINTS; PR01781; LATPROTEIN.
SQ SEQUENCE 242 AA; 26076 MW; D7A74C50CF5AB35 CRC64;

```

```

QY 60 GSHRTSPSRSDGANSVASYENEPAC--DADDEDYINPGYLVLPDSTPATSTRA 117
DB 93 GSHRMPSSQONSVDANSVASYENEPACKNDVDEVDYPR-GYLVLPDSSPAVPV 151
QY 118 PSAPALSTPGIRDSAFSMESIDYVNPESGESASASLDGSEYVNSOELHPGAKTEP 177
DB 152 SSAPVPSPNDLDSAFSPSCEDYVNPBESSESASASLDGSEYVNSPBPQP-VTRAE 210
QY 178 AALSSQEA-EVEEG---APDYENIQLIN 203
DB 211 ASVNSQVEDEBGECEGVDEAPDYENIQLIN 242

RESULT 7
Q9CAL8 PRELIMINARY; PRT; 710 AA.
ID Q9CAL8;
AC Q9CAL8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein F24J13.3.
GN Name=F24J13.3;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Roming C.M., Koo H., Fujii C.Y., Uterback T.R.,
RA Barnstead M.E., Bowman C.O., White O., Nierman W.C., Frazer C.M.;
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
CC -1-SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AC010796; AAG52479.1; -.
DR PIR; D96728; D96728.
DR HSP; P36897; IIAS.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot Kinase.
DR InterPro; IPR008271; Ser_Thr_kin_AS.
DR Pfam; PF00069; PKinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Hypothetical protein; kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 710 AA; 75371 MW; 8B4420A8D8F3A6 CRC64;

Query Match 12.3%; Score 132; DB 2; Length 710;
Best Local Similarity 31.6%; Pred. No. 3;
Matches 49; Conservative 6; Mismatches 52; Indels 48; Gaps 6;

```


OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxId=7955;
 RN (1)
 RC SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=whole body;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Schenker C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Locquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
 RA Boeck S.A., McEwan P.J., McKernan K.J., Malek J.A., Gamaralle P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN (12)
 RN SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=whole body;
 RA Struhsberg R.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC058052; AAHS052.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 754 AA; 84183 MW; D5DED29D4C1B7ADE CRC64;

Query Match 11.8%; Score 126.5; DB 2; Length 754;
 Best Local Similarity 26.3%; Pred. No. 7.1;
 Matches 50; Conservative 25; Mismatches 70; Indels 45; Gaps 10;

QY 16 PRGIQKRPHTVAPMPAPVPTSYRPLSQP-----DLI-----PIRSPQPLGSGSHRTP 65
 DB 422 PGGYL-----PPPPSPPLPTHTIPSPRAHHTLHGHSTYR-LLPQL-VRHDP 472
 QY 66 SSRSDGANSVASYENEPACEDADEDDYHNGYLVLPDSTPATSTAPAPALST 125
 DB 473 PSPPEPPAPRDEDDYDDEE-----EEDEDDDEDEPP-----PDHLPSQSQPLPRSR 523
 QY 126 PGIRSAASMSIDYVNVPPESGESASASLDGSRVYVNSQELHGAATK-PAALSSQ 184
 DB 524 CLVGLSLVSV-----IPEGNNSEEBED-----EEDQHPESDSDGFLVYKDE 567
 QY 185 AEEVEEGAP 194
 DB 568 SDEDEDDSP 577

RESULT 11
 O6LAT3 PRELIMINARY; PRT; 311 AA.
 ID O6LAT3
 AC O6LAT3
 DT 05-JUN-2004 (TrEMBLrel. 27, Created)
 DT 05-JUN-2004 (TrEMBLrel. 27, Last sequence update)
 DE 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)
 DR Hypothetical protein P0478F09.11.
 GN Name=P0478F09.11;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxId=39947;
 RN (1)

RP SEQUENCE FROM N.A.
 RA Chow T.-Y., Hsiang Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
 RA Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-X., Chen T.-R.,
 RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
 RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kuo P.-I., Lee M.-C.,
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
 RA Wu H.-P., Shaw J.-F.;
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC130610; AAT38083.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 311 AA; 33620 MW; 56CAF49DA2B3886B CRC64;

Query Match 11.5%; Score 123.5; DB 2; Length 311;
 Best Local Similarity 23.1%; Pred. No. 3.9;
 Matches 46; Conservative 30; Mismatches 84; Indels 39; Gaps 6;

QY 30 WPPAPVPTSYRPL-----SQPDLPIPRSPQPLGSGSHRPPSSRRSDGANSVASYENEP 86
 DB 53 WAPAPLITRSMADVEDDDDDYFATTAAPRVRWGTHHAADAHDDHDEQAALBQELSE 112
 QY 87 CEDADEDDYHNGYLVLPDSTPA-----TSTAAPAPALSTPG----- 127
 DB 113 DEVDVDDAEDHEH-----ETEDATPAEPANMKAAPAPPPOTERQSKKELKELEL 168
 QY 128 ---IRDSAFMSIDYVNVPPESGESASASLDGSRVYVNSQELHGAATKPEPALSSQ 184
 DB 169 DALIALELSSKSNNDACNETNGKKGAQADGE-----NKEGAPADESK---SSKK 218
 QY 185 AEEVEEGAPDENLELN 203
 DB 219 KKAKDKSAKAKETQELN 237

RESULT 12
 AAT38083 PRELIMINARY; PRT; 311 AA.
 ID AAT38083
 AC AAT38083
 DT 01-JUN-2004 (TrEMBLrel. 27, Created)
 DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
 DE 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
 DR Hypothetical protein P0478F09.11.
 GN P0478F09.11.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza; Oryza sativa.
 OC NCBI_TaxId=39947;
 RN (1)
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Chow T.-Y., Hsiang Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
 RA Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-X., Chen T.-R.,
 RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
 RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kuo P.-I., Lee M.-C.,
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
 RA Wu H.-P., Shaw J.-F.;
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC130610; AAT38083.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 311 AA; 33620 MW; 56CAF49DA2B3886B CRC64;

Query Match 11.5%; Score 123.5; DB 2; Length 311;
 Best Local Similarity 23.1%; Pred. No. 3.9;
 Matches 46; Conservative 30; Mismatches 84; Indels 39; Gaps 6;

QY 30 WPPAPVPTSYRPL-----SQPDLPIPRSPQPLGSGSHRPPSSRRSDGANSVASYENEP 86
 DB 53 WAPAPLITRSMADVEDDDDDYFATTAAPRVRWGTHHAADAHDDHDEQAALBQELSE 112
 QY 87 CEDADEDDYHNGYLVLPDSTPA-----TSTAAPAPALSTPG----- 127
 DB 113 DEVDVDDAEDHEH-----ETEDATPAEPANMKAAPAPPPOTERQSKKELKELEL 168

Oy	128----	IDSASFSMSIDYVNVPSGSEAEKSLDSREYVWVSOPHLHGCAKTEPALASOE	164
Dd	169	DALLAELELSSSKSNNDQNETNCKGAEGQAODE-----MKGAAPAPAESK-----SRKK	218
Oy	185	AEEVEEGADPYENIQELN	203
Dd	219	KKAKDKSAKEAKETQELN	237
 RESULT 13 NP14_RAT STANDARD; PRT; 704 AA.			
ID	NP14_RAT	PRT;	704 AA.
AC	P41777;		
DT	01-NOV-1995 (Rel. 32, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	Nucleolar phosphoprotein p130 (Nucleolar 130 kDa protein) (140 kDa nucleolar phosphoprotein) (Nopp140) (Nucleolar and coiled-body phosphoprotein 1).		
DE	Name=NCI:		
GN	Rattus norvegicus (Rat).		
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_Taxid:10116;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 292-309 AND 563-601.		
RC	TISSUE=Liver;		
RX	MEDLINE=922323542; PubMed=1623516;		
RX	Weier U.T., Blobel G.;		
RA	"Nopp140 shuttles on tracks between nucleolus and cytoplasm.";		
RL	Cell 70:127-138(1992).		
RN	[2]		
RP	INTERACTION WITH NOP5 AND FIBRILLARIN.		
RX	MEDLINE=20143579; PubMed=10679015.		
RA	Yang Y., Isaac C., Wang C., Dragoon F., Pogacic V., Weier U.T.;		
RT	"Conserved composition of mammalian box H/ACA and box C/D small nucleolar ribonucleoprotein particles and their interaction with the common factor Nopp140.";		
RL	Mol. Biol. Cell 11:567-577(2000).		
CC	-1- FUNCTION: Related to nucleogenesis, may play a role in the maintenance of the fundamental structure of the fibrillar center and dense fibrillar component in the nucleolus. It has intrinsic GTPase and ATPase activities. May play an important role in transcription catalyzed by RNA polymerase I (By similarity)".		
CC	-1- SUBUNIT: Interacts with DKC1/NAP57, NOP5/NAP65 and fibrillarin.		
CC	-1- SUBCELLULAR LOCATION: Shuttles on curvilinear tracks between nucleolus and cytoplasm. These tracks extend from the dense fibrillar component of the nucleolus across the nucleoplasm to a limited number of nuclear pore complexes.		
CC	-1- PTM: Undergoes rapid and massive phosphorylation/dephosphorylation cycles on CK2 and PKC sites. NOPP140 is one of the mostly phosphorylated proteins in the cell.		
CC	-1- SIMILARITY: Contains 1 Lish domain.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL, M94287; AAAA1718.1; -		
DR	EMBL, M94288; AAAA1719.1; -		
DR	PIR, B42680; B42680.		
DR	InterPro, IPR006594; Lish.		
DR	InterPro, IPR007718; SRP40_C.		
DR	InterPro, IPR003993; Treacle.		
DR	Pfam, PF05022; SRP40_C_1.		
DR	PRINTS, PR01503; TREACLE.		
DR	PROSITE, PS50896; LISH, 1.		
KW	ATP-binding; Direct protein sequencing; GTP-binding; Nuclear protein,		

KM	Phosphorylation; Repeat.
FT	DOMAIN 10 42
FT	DOMAIN 84 570
FT	REPEAT 84 95
FT	REPEAT 127 138
FT	REPEAT 170 181
FT	REPEAT 231 242
FT	REPEAT 274 285
FT	REPEAT 335 346
FT	REPEAT 373 384
FT	REPEAT 434 445
FT	REPEAT 479 490
FT	REPEAT 524 535
FT	REPEAT 559 570
FT	MOD_RES 567 567
FT	VARIANT 150 150
SQ	SEQUENCE 704 AA; 73562 MW; 1ADF1BF2DE483EA3_CRC64;
Query Match	11.4%; Score 122; DB 1; Length 704;
Beet Local Similarity	23.2%; Pred. No. 13;
Matches	54; Conservative 25; Mismatches 88; Indels 66; Gaps 6;
QY	5 SYDSTSSSLVPRGIQFKRPHTVAWPAPVPPTVSYPPLSQPDLPIP-----RSPQL 58
DB	372 SDDSDSDSSSDEDA-----PAKVVSKTSPLSKPATVPKPAKAIVATKKP 418
QY	59 GGSHTRSSRRDSOGANSVASYEENEPCADADEDEDYNNGYLVLPSTPATSTAAP 118
DB	419 AGSGGCKPGSRKADSSSEESSSEESBEATKKS-----VTPKAVATAKAAP 464
QY	119 SAPALSTR--GITDSAFMSISIDYVNP-----EGES 150
DB	465 SLPKQAQPARAGDSSPSSESSSEEEKTKPPKPAKKAAGAAPVKPTPVKKAASESSS 524
QY	151 AEASLDGSR-----YVNVSQELHPGAKTPEPALSSOEAEVEECPADYEN 198
DB	525 SSSSEDSSEEEKKPKSKATPKPPAQRANGVPASQNGKAGESEEBEDTEQN 577
RESULT 14	
GUPMO	PRELIMINARY; PRT; 253 AA.
ID	GUPMO
AC	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DR	01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE	Hypothetical protein DKFZp434H018 (Fragment).
GN	Name=DKFZp434H018;
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN	[1]
RP	NCBI_TaxId=9606;
RC	SEQUENCE FROM N.A.
RC	TISSUE=Testis;
RA	Ostenweider B., Obermaier B., Mewes H.W., Gaessenhuber J., Wiemann S.;
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AL117579; CAB56002.1; -
DR	PIR; T17312; T17312.
KW	Hypothetical protein.
FT	NON_TER 1
SQ	SEQUENCE 253 AA; 36911 MW; 36F210EA9318F791_CRC64;
Query Match	11.3%; Score 121.5; DB 2; Length 253;
Beet Local Similarity	23.6%; Pred. No. 4.1;
Matches	61; Conservative 32; Mismatches 63; Indels 103; Gaps 13;
QY	3 PGSYDS-----TSSDLVPRGIQFKRPHTVAWM-----APPV 37
DB	12 PGNSHSSTHSGRPEELPTRPEDPF-----PPLPLPWASHNRGGSPRCSPLSEAWPLT 64
QY	38 TSYPPLSQPDLPIP-RSFQPLGGSHTRPSSRRDSGANSVASYNEEPACADEDEDY 97

```
Db      65  TSSAPGCEPALPGPHEPSPGSGS---PALRREDLAAATPAS----- 103
Qy      98  HNPGLVLP-DSTPAT---TAAP-----SAPALSTPGIRDSAFSME----- 136
Db      104  ---PLVLPLETRPEEPQPSAPHPVPLSAPVPGSPDRKQSRSLIALSGLEKL 159
Qy      137  ----SIDDYVNVPESGESAB-----ASLDGSRVYVNSQELHPGAaktePALSSQ 183
Db      160  KTVTSGSIQPVQAQAGQGMVDTKRLKDSAVLDQSAKYTHLTHD-----ELISLLIQ 211
Qy      184  EAEVEEBGAPDYENLOEL 202
Db      212  RERELSQRD---EHVQEL 226

RESULT 15
Q9NEKO PRELIMINARY; PRT; 367 AA.
ID      Q9NEKO
AC      Q9NEKO;
DT      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Hypothetical protein Y116F11B.8.
GN      Name=Y116F11B.8;
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Peloderinae; Caenorhabditis.
OX      NCBI_taxid=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RX      MEDLINE=99069613; PubMed=9851916;
RA      none;
RT      "Genome sequence of the nematode C.elegans: A platform for
RT      investigating biology."
RL      Science 282:2012-2018 (1998).
[2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RA      Sulston J.E., McElay K.;
RA      Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AL132943; CAB81977.1; -.
DR      WormPep; Y116F11B.8; CE24165.
KW      Hypothetical protein.
SQ      SEQUENCE 367 AA; 40617 MW; DBEB05068359F0FB CRC64;

Query Match      11.3%; Score 121.5; DB 2; Length 367;
Best Local Similarity 25.9%; Pred. No. 6.3;
Matches 43; Conservative 18; Mismatches 68; Indels 37; Gaps 4;

Qy      36  PVTSTPPLSQPDLPIPRSPQPLGSGSHRTSSRRSDGANSVASYENEPACEDADEDED 95
Db      196  PVTSLPPPTDP-----PSTLGPMETPPNPGTSSAATTSAP----- 231
Qy      96  DYHNGYLVLPDSTPATSTAPASAPALSTPGIRDSAFSMEISIDYVNVPESGESASASL 155
Db      232  ----VTSBPATPEPASTLGPVVIWNWSPGTSSKATTSATVSLVLGPTTFSSAPQAL 283
Qy      156  DGSREYVNVSOELHPGAaktePALSSQAEVEE---GAPDYE 197
Db      284  DSSILA-ADPSTDPSPSISSTDPPTTSSTTAEVDYDEDDKGDDEYD 328
```

Search completed: November 21, 2004, 14:20:03
Job time : 197 secs

THIS PAGE BLANK (USPTO)